> Exhibit 13-12 Plot of MindWriter Scale Evaluation

What magazines do you want Singapore Airlines to carry for its in-flight service?

Most Preferred										Least Preferred	
10	9	8	7	6	5	4	3	2	1	0	(scale value)
3	4	7	10	13	16	13	10	7	4	3	(number of cards per pile)

> Cumulative Scales

Total scores on **cumulative scales** have the same meaning. Given a person's total score, it is possible to estimate which items were answered positively and negatively. A pioneering scale of this type was the scalogram. **Scalogram analysis** is a procedure for determining whether a set of items forms a unidimensional scale.[18] A scale is unidimensional if the responses fall into a pattern in which endorsement of the item reflecting the extreme position results in endorsing all items that are less extreme.

> **Exhibit 13-13** Ideal Scalogram Response Pattern*

*X = agree; — = disagree.

Assume we are surveying opinions regarding a new style of running shoe. We have developed a preference scale of four items:

1. The Airsole is good-looking.
2. I will insist on Airsole next time because it is great-looking.
3. The appearance of Airsole is acceptable to me.
4. I prefer the Airsole style to other styles.

Participants indicate whether they agree or disagree. If these items form a unidimensional scale, the response patterns will approach the ideal configuration shown in Exhibit 13-13. Item 2 is the most extreme position of the four attitude statements. A participant who agrees with item 2 will agree with all four items. The items are ordered in the scalogram left to right from most to least extreme. If each agreement renders a score of 1, a score of 4 indicates all statements are agreed upon and represents the most favorable attitude. Persons with a score of 3 should disagree with item 2 but agree with all others, and so on. According to scalogram theory, this pattern confirms that the universe of content (attitude toward the appearance of this running shoe) is scalable.

The scalogram and similar procedures for discovering underlying structure are useful for assessing attitudes and behaviors that are highly structured, such as social distance, organizational hierarchies, and evolutionary product stages.[19] The scalogram is used much less often today, but retains potential for specific applications.

>summary

1 Managers know that the measurement of attitudes is an important aspect of strategy and often the best tool available because attitudes reflect past experience and shape future behavior. Attitudes are learned, stable predispositions to respond to oneself, other persons, objects, or issues in a consistently favorable or unfavorable way. Attitudes are generally thought to be composed of three components: affective, cognitive, and behavioral intentions.

2 Selecting and constructing a measurement scale requires the consideration of several factors that influence the reliability, validity, and practicality of the scale. Two broad research objectives are to measure characteristics of the individuals who participate in studies and to use participants as judges of the objects or indicants presented to them. Measurement scales fall into one of four general response types: rating, ranking, categorization, and sorting. The properties of data are classified in increasing order of power—nominal, ordinal, interval, or ratio—which determines how a particular measurement scale's data will be analyzed statistically..Measurement scales are either unidimensional or multidimensional. A balanced rating scale has an equal number of categories above and below the midpoint, whereas an unbalanced rating scale has an unequal number of favorable and unfavorable response choices. An unforced-choice rating scale provides participants with an opportunity to express no opinion when they are unable to make a choice among the alternatives offered. A forced-

choice scale requires that they select one of the offered alternatives. The ideal number of points for a rating scale should match the stimulus presented and extract information proportionate to the complexity of the attitude object. The value of rating scales depends on the assumption that a rater can and will make good judgments. Errors of central tendency, halo effect, and leniency adversely affect a precise understanding of the measurement.

3 Rating scales have several uses, design features, and requirements. The simple category scale offers two mutually exclusive response choices. The multiple-choice, single-response scale offers the rater several options, including "other." The multiple-choice, multiple-response scale (also called a checklist) allows the rater to select one or several alternatives, thereby providing a cumulative feature.

The Likert scale consists of a series of statements, and the participant is asked to agree or disagree with each statement. Summation is possible with this scale although *not* necessary and in some instances undesirable.

The semantic differential (SD) scale measures the psychological meanings of an attitude object. Researchers use this scale for studies of brand and institutional image. The method consists of a set of bipolar rating scales, usually with 7 points, by which one or more participants rate one or more concepts on each scale item. The Stapel scale is used as an alternative to the semantic differential, especially when it is difficult to find bipolar adjectives that match the investigative question. Participants select a plus number for the characteristic that describes the attitude object. Ratings range from +5 to −5, where participants select a number that describes the object very accurately to very inaccurately.

Numerical scales have equal intervals that separate their numeric scale points. Verbal anchors serve as the labels for the extreme points. Numerical scales are often 5-point scales but may have 7 or 10 points.

A multiple rating list scale is similar to the numerical scale but accepts a circled response from the rater, and the layout allows visualization of the results.

A scale that helps the researcher discover proportions is the constant-sum scale. The participant distributes 100 points among up to 10 categories. The graphic rating scale was originally created to enable researchers to discern fine differences. Raters check their response at any point along a continuum. Other graphic rating scales use pictures, icons, or other visuals to communicate with children or others whose limited vocabulary prevents the use of scales anchored with words.

Ranking scales allow the participant to compare two or more objects and make choices among them. Frequently, the participant is asked to select one as the "best" or the "most preferred." When there are only two choices, as with the paired-comparison scale, the participant can express attitudes unambiguously by choosing between two objects. The forced ranking scale lists attributes that are ranked relative to each other. This method is faster than paired comparisons and more user-friendly. Often the researcher is interested in benchmarking. This calls for a standard by which training programs, processes, brands, point-of-sale purchases, or people can be compared. The comparative scale is ideal for such comparisons if the participants are familiar with the standard.

Q-sorts are a form of scaling that requires sorting of a deck of cards into piles that represent points along a continuum. The purpose of sorting is to get a conceptual representation of the sorter's attitude toward the attitude object and to compare the relationships between people. Given a person's total score, it is possible to estimate which items were answered positively and negatively on cumulative scales. A pioneering cumulative scale was the scalogram, a procedure for determining whether a set of items forms a unidimensional scale.

>keyterms

>discussionquestions

Terms in Review

1 Discuss the relative merits of and problems with:
   a Rating and ranking scales.
   b Likert and differential scales.
   c Unidimensional and multidimensional scales.

Making Research Decisions

2 Suppose your firm had planned a major research study for November 2001. Given the incidents of September 11, your superior decides to add a question to the study. The question must measure consumers' confidence that the U.S. economic system will be able to rebound following the terrorist attacks of September and the subsequent effects of those incidents (increased layoffs, higher unemployment, numerous firms failing to meet their sales and profit projections, lower holiday retail sales, war on terrorism). Draft a scale of each of the following types to measure that confidence level:
   a Constant-sum scale.
   b Likert-type summated scale.
   c Semantic differential scale.
   d Stapel scale.
   e Forced ranking scale.

3 An investigative question in your employee satisfaction study seeks to assess employee "job involvement." Create a measurement question that uses the following scales:
   a A graphic rating scale.
   b A multiple rating list.
   Which scale do you recommend and why?

4 You receive the results of a paired-comparison preference test of four soft drinks from a sample of 200 persons. The results are as follows:

Mr. Peepers	Koak	Zip	Pabze
Koak	—	50*	115
35			
Zip	150	—	160
70			
Pabze	85	40	—
45			
Mr. Peepers	165	130	155
—			

*Read as 50 persons preferred Zip to Koak.

a How do these brands rank in overall preference in this sample?
b Develop an interval scale for these four brands.

5 One of the problems in developing rating scales is the choice of response terms to use. Below are samples of some widely used scaling codes. Do you see any problems with them?
   a Yes—Depends—No
   b Excellent—Good—Fair—Poor
   c Excellent—Good—Average—Fair—Poor
   d Strongly Approve—Approve—Uncertain—Disapprove—Strongly Disapprove

6 You are working on a consumer perception study of four brands of bicycles. You will need to develop measurement questions and scales to accomplish the tasks listed below. Be sure to explain which data levels (nominal, ordinal, interval, ratio) are appropriate and which quantitative techniques you will use.
   a Prepare an overall assessment of all the brands.
   b Provide a comparison of the brands for each of the following dimensions:
   (1) Styling
   (2) Durability
   (3) Gear quality
   (4) Brand image

7 Below is a Likert-type scale that might be used to evaluate your opinion of the educational degree program in which you are enrolled. There are five response categories: Strongly Agree, Agree, Neither Agree nor Disagree, Disagree, and Strongly Disagree. If Strongly Agree (SA) represents the most positive attitude, how would you value the items below? Record your answers to the items.
   a This program is not very challenging.   SA A N D SD
   b The general level of teaching is good.   SA A N D SD
   c I really think I am learning a lot from this program.   SA A N D SD
   d Students' suggestions are given little attention here.   SA A N D SD
   e This program does a good job of preparing one for a career.   SA A N D SD

f This program is below my expectations. SA A N D SD

In what two different ways could such responses be used? What would be the purpose of each?

Bringing Research to Life

8 What is the basis of Jason and Myra's argument for the need of an arbitrary scale to address customer expectations?

From Concept to Practice

9 Using the response strategies within Exhibit 13-1 or 13-9, which would be appropriate and add insight to understanding the various indicants of student demand for the academic program in which the students are enrolled?

>wwwexercise

Visit the Websurveyor.com site and review the sample surveys. In the Software Use survey, evaluate the question "How important are the following items in making a decision to purchase a software product?" Notice that Websurveyor uses a 3-point scale (very important, important, not important) to illustrate a matrix question. What are the advantages and disadvantages of this design?

>cases*

BBQ Product Crosses over the Lines of Varied Tastes	**Mastering Teacher Leadership**
Calling Up Attendance	**NCRCC: Teeing Up and New Strategic Direction**
Campbell-Ewald: R-E-S-P-E-C-T Spells Loyalty	**Ramada Demonstrates Its Personal Best™**
Donatos: Finding the New Pizza	**USTA: Come Out Swinging**
Inquiring Minds Want to Know—NOW!	**Yahoo!: Consumer Direct Marries Purchase Metrics to Banner Ads**

* All cases appear on the text CD; you will find abstracts of these cases in the Case Abstracts section of this text. Video cases are indicated with a video icon.

>chapter 14

Questionnaires and Instruments

❝ ❝By using the Internet, you can show consumers pictures, show them packaging and even play videos.❞ ❞

Gordon Black, founder, Harris Interactive

>learningobjectives

After reading this chapter, you should understand . . .

1 The link forged between the management dilemma and the communication instrument by the management-research question hierarchy.

2 The influence of the communication method on instrument design.

3 The three general classes of information and what each contributes to the instrument.

4 The influence of question content, question wording, response strategy, and preliminary analysis planning on question construction.

5 Each of the numerous question design issues influencing instrument quality, reliability, and validity.

6 Sources for measurement questions.

7 The importance of pretesting questions and instruments.

"How is the Albany questionnaire coming?" asks Jason as he enters Sally's office.

"The client approved the investigative questions this morning. So we are ready to choose the measurement questions and then write the questionnaire," shares Sally, glancing up from her computer screen. "I was just checking our bank of pretested questions. I'm looking for questions related to customer satisfaction in the medical field."

"If you are already searching for appropriate questions, you must have the analysis plan drafted. Let me see the dummy tables you developed," requests Jason. "I'll look them over while you're scanning."

Sally hands over a sheaf of pages. Each has one or more tables referencing the desired information variables. Each table indicates the statistical diagnostics that would be needed to generate the table.

As the computer finishes processing, Sally scans the revealed questions for appropriate matches to Albany's information needs. "At first glance, it looks like there are several multiple-choice scales and ranking questions we might use. But I'm not seeing a rating scale for overall satisfaction. We may need to customize a question just for Albany."

"Custom designing a question is expensive. Before you make that choice," offers Jason, "run another query using *CardioQuest* as a keyword. A few years ago, I did a study for that large cardiology specialty in Orlando. I'm sure it included an overall satisfaction scale. It might be worth considering."

Sally types *CardioQuest* and *satisfaction,* and then waits for the computer to process her request. "Sure enough, he's right again," murmurs Sally. "How do you remember all the details of prior studies done eons ago?" she asks, throwing the purely hypothetical question at Jason. But Sally swivels to face Jason, all senses alert when she hears his muffled groan.

Jason frowns as he comments, "You have far more analytical diagnostics planned than would be standard for a project of this type and size, Sally. For example, are Tables 2, 7, and 10 really necessary?" Jason pauses but doesn't allow time for Sally to answer. "To stay within budget, we are going to have to whittle down the analysis phase of the project to what is essential. Let's see if we can reduce the analysis plan to something that we both can live with. Now, walk me through what you think you'll reveal by three-way cross-tabulating these two attitudinal variables with the education variable."

New researchers often want to draft questions immediately. Their enthusiasm makes them reluctant to go through the preliminaries that make for successful surveys. Exhibit 14-1 is a suggested flowchart for instrument design. The procedures followed in developing an instrument vary from study to study, but the flowchart suggests three phases. Each phase is discussed in this chapter, starting with a review of the research question hierarchy.

> Revisiting the Research Question Hierarchy

The management-research question hierarchy is the foundation of the research process and also of successful instrument development (see Exhibit 14-2). By this stage in a research project, the process of moving from the general management dilemma to specific measurement questions has traveled through the first three question levels:

This section relates to
Phase 1 in Exhibit 14-1.

1. *Management question*—the dilemma, stated in question form, that the manager needs resolved.

2. *Research question(s)*—the fact-based translation of the question the researcher must answer to contribute to the solution of the management question.

3. *Investigative questions*—specific questions the researcher must answer to provide sufficient detail and coverage of the research question. Within this level, there may be several questions as the researcher moves from the general to the specific.

4. *Measurement questions*—questions participants must answer if the researcher is to gather the needed information and resolve the management question.

< We discussed how to refine a management dilemma and take it through the research process in Chapter 3 and depicted the process in Exhibit 3-2.

In the Albany Outpatient Laser Clinic study, the eye surgeons would know from experience the types of medical complications that could result in poor recovery. But they might be far less knowledgeable about what medical staff actions and attitudes affect client recovery and perception of well-being. Coming up with an appropriate set of information needs in this study will take the guided expertise of the researcher. Significant exploration would likely have preceded the development of the investigative questions. In the project for MindWriter, exploration was limited to several interviews and data mining of

> Exhibit 14-1 Overall Flowchart for Instrument Design
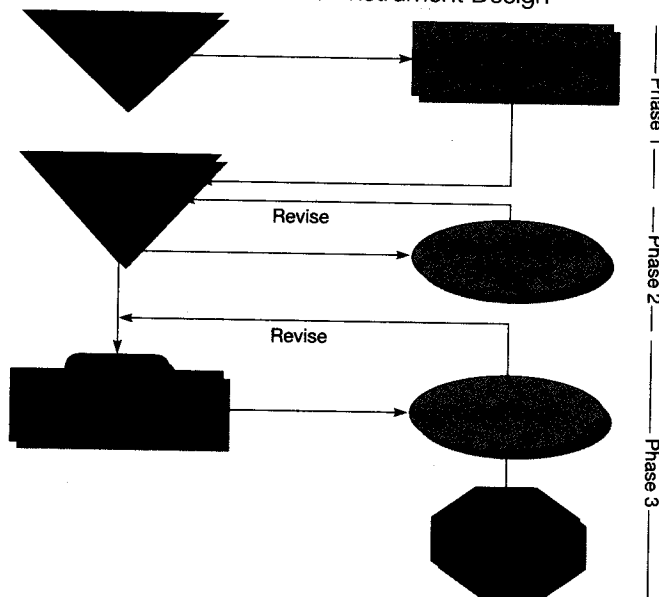
company service records because the concepts were not complicated and the researchers had experience in the industry.

Normally, once the researcher understands the connection between the investigative questions and the potential measurement questions, a strategy for the survey is the next logical step. This proceeds to getting down to the particulars of instrument design. The following are prominent among the strategic concerns:

> **The Closeup at the end of this chapter reveals the thinking that led to the final questionnaire in the MindWriter CompleteCare project.**

1. What type of scale is needed to perform the desired analysis to answer the management question?
2. What communication approach will be used?
3. Should the questions be structured, unstructured, or some combination?
4. Should the questioning be undisguised or disguised? If the latter, to what degree?

Technology has also affected the survey development process, not just the method of the survey's delivery. Today's software, hardware, and Internet and intranet infrastructures allow researchers to (1) write questionnaires more quickly by tapping question banks for appropriate, tested questions, (2) create visually driven instruments that enhance the process for the participant, (3) use questionnaire software that eliminates separate manual data entry, and (4) build questionnaires that save time in data analysis.[1]

Type of Scale for Desired Analysis

The analytical procedures available to the researcher are determined by the scale types used in the survey. As Exhibit 14-2 clearly shows, it is important to plan the analysis before developing the measurement questions. Chapter 12 discussed nominal, ordinal, interval, and ratio scales and explained how the characteristics of each type influence the analysis (statistical choices and hypothesis testing). We demonstrate how to code and extract the data from the instrument, select appropriate descriptive measures or tests, and analyze the results in Chapters 16 to 20. In this chapter, we are most interested in asking each question in the right way and in the right order to collect the appropriate data for desired analysis.

Communication Approach

As discussed in Chapter 10, communication-based research may be conducted by personal interview, telephone, mail, computer (intranet and Internet), or some combination of these (called *hybrid studies*). Decisions

> **Exhibit 14-2** Flowchart for Instrument Design: Phase 1
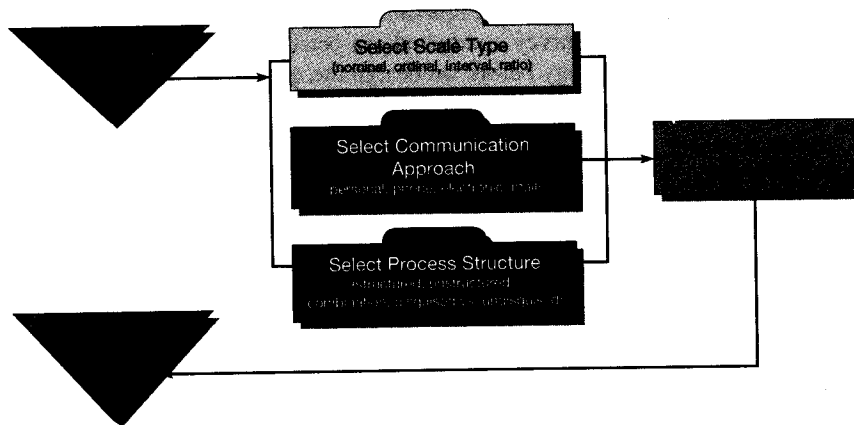
WebSurveyor is one of the earliest developed and one of the most used desktop survey design software packages. WebSurveyor 5.0, a new high-end survey tool that
helps organizations rapidly gather and analyze mission-critical data, is offered free to qualifying universities and colleges through the WebSurveyor Academic Grant program (**www.websurveyor.com/ grant**). You can take a tour of this powerful survey design and reporting tool at **www.websurveyor.com.**
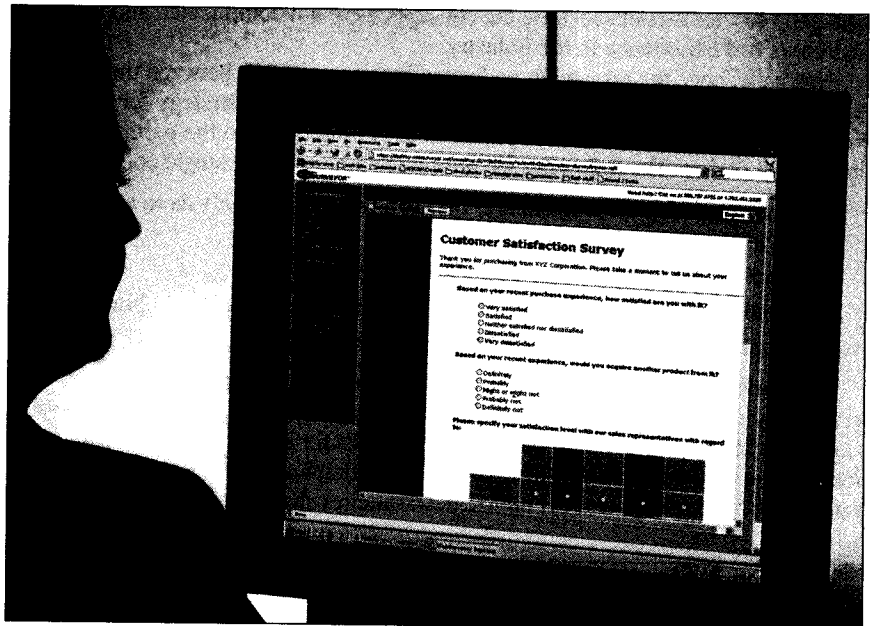
You'll find tips for intercept questionnaire design on your text CD.

regarding which method to use as well as where to interact with the participant (at home, at a neutral site, at the sponsor's place of business, etc.) will affect the design of the instrument. In personal interviewing and computer surveying, it is possible to use graphics and other questioning tools more easily than it is in questioning done by mail or phone. The different delivery mechanisms result in different introductions, instructions, instrument layout, and conclusions. For example, researchers may use intercept designs, conducting personal interviews with participants at central locations like shopping malls, stores, sports stadiums, amusement parks, or county fairs. The intercept study poses several instrument challenges.

In the MindWriter example, these decisions were easy. The dispersion of participants, the necessity of a service experience, and budget limitations all dictated a mail survey where the participant received the instrument either at home or at work. Using a telephone survey, which in this instance is the only way to follow up with nonparticipants, could, however, be problematic. This is due to memory decay caused by the passage of time between return of the laptop and contact with the participant by telephone.

Jason and Sally have several options for the Albany study. Clearly a self-administered study is possible, as all the participants are congregating in a centralized location for scheduled surgery. But given the importance of some of the information to medical recovery, a survey conducted via personal interview might be an equally valid choice. We need to know the methodology before we design the questionnaire, as some measurement scales are difficult to answer without the visual aid of seeing the scale.

Disguising Objectives and Sponsors

Another consideration in communication instrument design is whether the purpose of the study should be disguised. A **disguised question** is designed to conceal the question's true purpose. Some degree of disguise is often present in survey questions, especially to shield the study's sponsor. We disguise the sponsor and the objective of a study if the researcher believes that participants will respond differently than they would if both or either was known.

The accepted wisdom among researchers is that they must disguise the study's objective or sponsor in order to obtain unbiased data. The decision about when to use disguised questions within surveys may be made easier by identifying four situations where disguising the study objective is or is not an issue:

- Willingly shared, conscious-level information.
- Reluctantly shared, conscious-level information.
- Knowable, limited-conscious-level information.
- Subconscious-level information.

In surveys requesting conscious-level information that should be willingly shared, either disguised or undisguised questions may be used, but the situation rarely requires disguised techniques.

Example: Have you attended the showing of a foreign language film in the last six months?

In the MindWriter study, the questions revealed in Exhibit 14-12 ask for information that the participant should know and be willing to provide.

< **You might wish to review the projective techniques discussed in Chapter 8.**

Sometimes the participant knows the information we seek but is reluctant to share it for a variety of reasons. When we ask for an opinion on some topic on which participants may hold a socially unacceptable view, we often use projective techniques. In this type of disguised question, the survey designer phrases the questions in a hypothetical way or asks how other people in the participant's experience would answer the question. We use projective techniques so that participants will express their true feelings and avoid giving stereotyped answers. The assumption is that responses to these questions will indirectly reveal the participants' opinions.

Example: Have you downloaded copyrighted music from the Internet without paying for it? (nonprojective)

Example: Do you know people who have downloaded copyrighted music from the Internet without paying for it? (projective)

Not all information is at the participant's conscious level. Given some time—and motivation—the participant can express this information. Asking about individual attitudes when participants know they hold the attitude but have not explored why they hold the attitude may encourage the use of disguised questions. A classic example is a study of government bond buying during World War II.[2] A survey sought reasons why, among people with equal ability to buy, some bought more war bonds than others. Frequent buyers had been personally solicited to buy bonds, while most infrequent buyers had not received personal solicitation. No direct *why* question to participants could have provided the answer to this question because participants did not know they were receiving differing solicitation approaches.

Example: What is it about air travel during stormy weather that attracts you?

In assessing buying behavior, we accept that some motivations are subconscious. This is true for attitudinal information as well. Seeking insight into the basic motivations underlying attitudes or consumption practices may or may not require disguised techniques. Projective techniques (such as sentence completion tests, cartoon or balloon tests, and word association tests) thoroughly disguise the study objective, but they are often difficult to interpret.

Example: Would you say, then, that the comment you just made indicates you would or would not be likely to shop at Galaxy Stores? (survey probe during personal interview)

In the MindWriter study, the questions were direct and undisguised, as the specific information sought was at the conscious level. Customers knew they were evaluating their experience with the service and repair program at MindWriter; thus the purpose of the study and its sponsorship were also

> **The MindWriter questionnaire is Exhibit 14-12, p. 384.**

> Exhibit 14-3 Dummy Table for American Eating Habits

Age	**Use of Convenience Foods**				
	Always Use	Use Frequently	Use Sometimes	Rarely Use	Never Use
18-24					
25-34					
35-44					
45-54					
55-64					
65+					

undisguised. While the sponsor of the Albany Clinic study was obvious, any attempt by a survey to reveal psychological factors that might affect recovery and satisfaction might need to use disguised questions. The survey would not want to unnecessarily upset a patient before or immediately following surgery, as that might in itself affect attitude and recovery.

Preliminary Analysis Plan

Researchers are concerned with adequate coverage of the topic and with securing the information in its most usable form. A good way to test how well the study plan meets those needs is to develop "dummy" tables that display the data one expects to secure. Each **dummy table** is a cross-tabulation between two or more variables. For example, in the biennial study of what Americans eat conducted by *Parade* magazine,[3] we might be interested to know whether age influences the use of convenience foods. The dummy table shown in Exhibit 14-3 would match the age ranges of participants with the degree to which they use convenience foods. The preliminary analysis plan serves as a check on whether the planned measurement questions (for example, the rating scales on use of convenience foods and on age) meet the data needs of the research question. This also helps the researcher determine the type of scale needed for each question (for example, ordinal data on frequency of use and on age)—a preliminary step to developing measurement questions for investigative questions.

< You might find it useful to review Exhibit 10-1, "Data Collection Approach," in Chapter 10.

In the opening vignette, Jason and Sally use the development of a preliminary analysis plan to determine whether the project could be kept on budget. The number of hours spent on data analysis is a major cost of any survey. Too expansive an analysis plan can reveal unnecessary questions. *The guiding principle of survey design is always to ask only what is needed.*

> Constructing and Refining the Measurement Questions

This section relates to Phase 2 in Exhibit 14-1.

Drafting or selecting questions begins once you develop a complete list of investigative questions and decide on the collection processes to be used. The creation of a survey question is not a haphazard or arbitrary process. It is exacting and requires paying significant attention to detail and simultaneously addressing numerous issues. Whether you create or borrow or license a question, in Phase 2 (see Exhibit 14-4) you generate specific measurement questions considering subject content, the wording of each question (influenced

> **Exhibit 14-4** Flowchart for Instrument Design: Phase 2
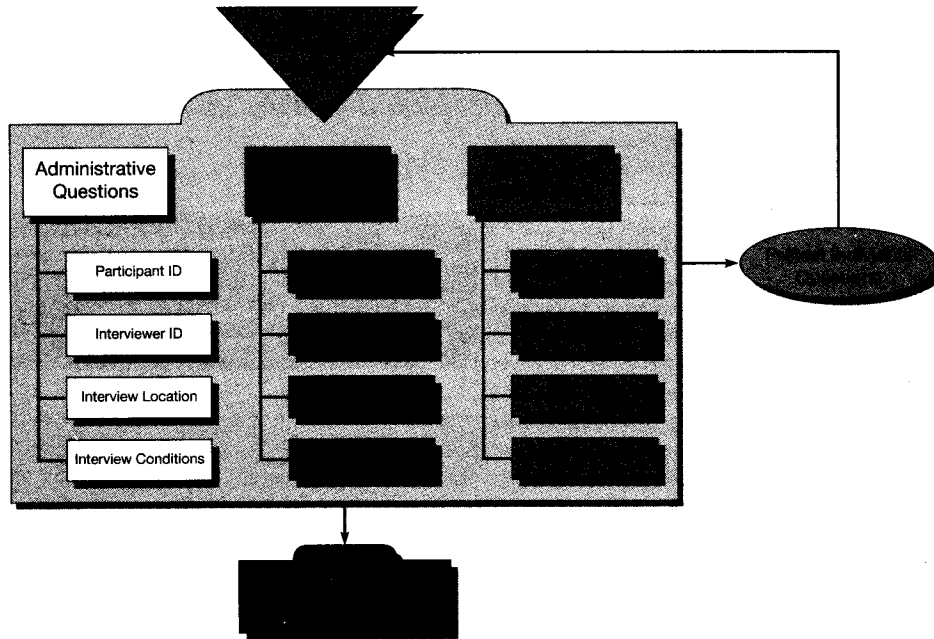
by the degree of disguise and the need to provide operational definitions for constructs and concepts), and response strategy (each producing a different level of data as needed for your preliminary analysis plan). In Phase 3 you must address topic and question sequencing. We discuss these topics sequentially, although in practice the process is not linear. For this discussion, we assume the questions are structured.

The order, type, and wording of the measurement questions, the introduction, the instructions, the transitions, and the closure in a quality communication instrument should accomplish the following:

• Encourage each participant to provide accurate responses.

• Encourage each participant to provide an adequate amount of information.

• Discourage each participant from refusing to answer specific questions.

• Discourage each participant from early discontinuation of participation.

• Leave the participant with a positive attitude about survey participation.

Question Categories and Structure

Questionnaires and **interview schedules** (an alternative term for the questionnaires used in personal interviews) can range from those that have a great deal of structure to those that are essentially unstructured. Questionnaires contain three categories of measurement questions:

• Administrative questions.

• Classification questions.

• Target questions (structured or unstructured).

>snapshot

A Survey Cold as Ice

In December 2002, 1.5 million North Carolinians lost power due to a storm that covered the state with more than an inch of ice that coated tree limbs and brought down power lines. After seven years of repeated natural disasters (two hurricanes, a record-setting flood, a major snowstorm, and drought), North Carolina needed to assess the true cost of this latest disaster and see how the state's residents used weather predictions to prepare for it. It is the state's responsibility to determine the effectiveness of its communication initiatives. The state decided to use a survey.

Odum Institute at the University of North Carolina–Chapel Hill, which maintains one of the nation's largest archives of polling data, and RTI International (a nonprofit "dedicated to conducting research that improves the human condition") stepped in. They conducted a telephone survey of 457 households in 36 counties—those counties included in North Carolina's application for federal disaster assistance. The goal was to give decision makers and the public information that would be useful in preparing for the state's next natural disaster.

RTI knew its analysis plan needed to reveal not only the direct costs of dealing with the disaster but also the indirect costs, like missed work, damage to residences, spoiled food, or hotel accommodations. It also knew that residents' satisfaction with emergency response would influence their actions during the next disaster. And RTI needed to measure attitudes about prevention. Would willingness to prevent a reoccurrence of lost power by burying power lines increase with number of days of power lost?

Shown in the right column is one dummy table for the RTI study for North Carolina, including the aggregate data that were collected. What types of measurement scales would be necessary to complete this table?

> Overall Satisfaction

	Power Supplier Response Satisfaction Rating		
Days without Power	Municipal Power	Duke Power	Progress Energy
	7.6	6.6	6.5

Eighty percent of households indicated a willingness to take preventive actions—including 47 percent that were willing to pay extra on their monthly bill to bury power lines. But one of the more significant findings was that the municipal power companies responded more quickly and earned a higher performance rating from customers than did Duke Power and Progress Energy. On a 10-point satisfaction scale ("I was satisfied with my electric power company's response to the ice storm") where 5 was "neither agree nor disagree," municipals earned a 7.6 while the nonmunicipal electric providers earned 6.6 and 6.5, respectively. Each day without power led to a decline in the household's satisfaction level—and the governor and local politicians suffered a similar fate. This survey had a margin of error of ±4.7 percent.

www.rti.com

Administrative questions identify the participant, interviewer, interview location, and conditions. These questions are rarely asked of the participant but are necessary for studying patterns within the data and identify possible error sources. **Classification questions** usually cover sociological-demographic variables that allow participants' answers to be grouped so that patterns are revealed and can be studied. These questions usually appear at the end of a survey (except for those used as *filters* or *screens,* questions that determine whether a participant has the requisite level of knowledge to participate). **Target questions** address the investigative questions of a specific study. These are grouped by topic in the survey. Target questions may be **structured** (they present the participants with a fixed set of choices; often called *closed questions*) or **unstructured** (they do not limit responses but do provide a frame of reference for participants' answers; sometimes referred to as *open-ended questions*).

At the Albany Clinic, some questions will need to be unstructured because anticipating medications and health history for a wide variety of individuals would be a gargantuan task for a researcher and would take up far too much paper space.

Question Content

Question content is first and foremost dictated by the investigative questions guiding the study. From these questions, questionnaire designers craft or borrow the target and classification questions that will be asked of participants. Four questions, covering numerous issues, guide the instrument designer in selecting appropriate question content:

- Should this question be asked (does it match the study objective)?
- Is the question of proper scope and coverage?
- Can the participant adequately answer this question as asked?
- Will the participant willingly answer this question as asked?

Exhibit 14-5 summarizes these issues related to constructing and refining measurement questions that are described below and detailed in Appendix 14a: "Crafting Effective Measurement Questions."

Question Wording

It is frustrating when people misunderstand a question that has been painstakingly written. This problem is partially due to the lack of a shared vocabulary. The difficulty of understanding long and complex sentences or involved phraseology aggravates the problem further. Our dilemma arises from the requirements of question design (the need to be explicit, to present alternatives, and to explain meanings). All contribute to longer and more involved sentences.[4]

The difficulties caused by question wording exceed most other sources of distortion in surveys. They have led one social scientist to conclude:

To many who worked in the Research Branch it soon became evident that error or bias attributable to sampling and to methods of questionnaire administration were relatively small as compared with other types of variations—especially variation attributable to different ways of wording questions.[5]

While it is impossible to say which wording of a question is best, we can point out several areas that cause participant confusion and measurement error. The diligent question designer will put a survey question through many revisions before it satisfies these criteria:[6]

- Is the question stated in terms of a shared vocabulary?
- Does the question contain vocabulary with a single meaning?
- Does the question contain unsupported or misleading assumptions?
- Does the question contain biased wording?
- Is the question correctly personalized?
- Are adequate alternatives presented within the question?

In the vignette, Sally's study of the prior survey used by the Albany Laser Clinic illustrated several of these problems. One question asked participants to identify their "referring physician" and the "physician most knowledgeable about your health." This question was followed by one requesting a single phone number. Participants didn't know which doctor's phone number was being requested. By offering space for only one number, the data collection instrument implied that both parts of the question might refer to the same doctor. Further, the questions about past medical history did not offer clear directions. One question asked participants about whether they had "had the flu recently," yet made no attempt to define whether *recently* was within the last 10 days or the last year. Another asked "Are your teeth intact?" Prior participants had answered by providing information about whether they wore false teeth, had loose teeth, or had broken or chipped teeth—only one of which was of interest to the doctor performing surgery. To another question ("Do you have limited motion of your neck?"), all

> **Exhibit 14-5** A Summary of the Major Issues Related to Measurement Questions

Issue Category	Fundamental Issue
Question Content	
Question Wording	
Response Strategy Choice	

respondents answered yes. Sally could only conclude that a talented researcher did not design the clinic's previously used questionnaire. While the Albany Outpatient Laser Clinic survey did not reveal any **leading questions,** these can inject significant error by inferring that one response should be favored over another. One classic hair care study asked, "How did you like Brand X when it lathered up so nicely?" Obviously, the participant was supposed to factor in the richness of the lather in evaluating the shampoo.

The MindWriter questionnaire (see Exhibit 14-12) simplified the process by using the same response strategy for each factor the participant was asked to evaluate. The study basically asks, "How did our CompleteCare service program work for you when you consider each of the following factors?" It accomplishes this by setting up the questioning with "Take a moment to tell us how well we've served you." Because the sample includes CompleteCare users only, the underlying assumption that participants have used the service is acceptable. The language is appropriate for the participant's likely level of education. And the open-ended question used for "comments" adds flexibility to capture any unusual circumstances not covered by the structured list.

Target questions need not be constructed solely of words. Computer-assisted, computer-administered, and Web surveys and interview schedules, and to a lesser extent printed surveys, often incorporate visual images as part of the questioning process.

Response Strategy

A third major decision area in question design is the degree and form of structure imposed on the participant. The various response strategies offer options that include **unstructured response** (or *open-ended response,* the free choice of words) and **structured response** (or *closed response,* specified alternatives provided). Free responses, in turn, range from those in which the participants express themselves extensively to those in which participants' latitude is restricted by space, layout, or instructions to choose one word or phrase, as in a fill-in question. Closed responses typically are categorized as dichotomous, multiple-choice, checklist, rating, or ranking response strategies.

Several situational factors affect the decision of whether to use open-ended or closed questions.[7] The decision is also affected by the degree to which these factors are known to the interviewer. The factors are:

- Objectives of the study.
- Participant's level of information about the topic.
- Degree to which participant has thought through the topic.
- Ease with which participant communicates.
- Participant's motivation level to share information.

All of the strategies that are described below are available for use on Web questionnaires. However, with the Web survey you are faced with slightly different layout options for response, as noted in Exhibit 14-6. For the multiple-choice or dichotomous response strategies, the designer chooses between radio buttons and drop-down boxes. For the checklist or multiple response strategy, the designer must use the checkbox. For rating scales, designers may use pop-up windows that contain the scale and instructions, but the response option is usually the radio button. For ranking questions, designers use radio buttons, drop-down boxes, and textboxes. For the free response question, the designer chooses either the one-line textbox or the scrolled textbox. Web surveys and other computer-assisted surveys can return participants to a given question or prompt them to complete a response when they click the "submit" button; this is especially valuable for checklists, rating scales, and ranking questions.

Free-Response Question

Free-response questions, also known as *open-ended questions,* ask the participant a question and either the interviewer pauses for the answer (which is

< You may wish to review Exhibits 13-2 and 13-9. These provide other question samples.

> **Exhibit 14-6** Internet Survey Response Options

unaided) or the participant records his or her ideas in his or her own words in the space provided on a questionnaire. Survey researchers usually try to reduce the number of such questions as they pose significant problems in interpretation and are costly in terms of data analysis.

Dichotomous Question

A topic may present clearly dichotomous choices: Something is a fact or it is not; a participant can either recall or not recall information; a participant attended or didn't attend an event. **Dichotomous questions** suggest opposing responses, but this is not always the case. One response may be so unlikely that it would be better to adopt the middle-ground alternative as one of the two choices. For example, if we ask participants whether a product is underpriced or overpriced, we are not likely to get many selections of the former choice. The better alternatives to present to the participant might be "fairly priced" or "overpriced."

In many two-way questions, there are potential alternatives beyond the stated two alternatives. If the participant cannot accept either alternative in a dichotomous question, he or she may convert the question to a multiple-choice or rating question by writing in his or her desired alternative. For example, the participant may prefer an alternative such as "don't know" to a yes-no question or prefer "no opinion" when faced with a favor-oppose option. In other cases, when there are two opposing or complementary choices, the participant may prefer a qualified choice ("yes, if X doesn't occur," or "sometimes yes and sometimes no," or "about the

> **Exhibit 14-6** (Cont'd)

Please select your answer

PC Magazine
Wired
Computing Magazine
Computing World
PC Computing
Laptop

Multiple Choice, Single Response
using pull–down box

Checklist
using checkbox
(may also use radio buttons)

Which of the following computing magazines did you look at in the last 60 days?

☐ PC Magazine
☑ Wired
☐ Computing Magazine
☐ Computing World
☐ PC Computing
☐ Laptop

Rating Grid
(may also use checkboxes)
Requires a single response per line
The longer the list, the more likely
the participant must scroll.

Fast reliable repair service	◉	○	○	○	○
Service at my location	○	○	◉	○	○
Maintenance by the manufacturer	◉	○	○	○	○
Knowledgeable technicians	◉	○	○	○	○
Notification of upgrades	○	○	○	○	◉

Ranking Question
using pull-down box
(may also use textboxes,
in which ranks are entered)
[This question asks for
a limited ranking of
only three of the
listed elements.]

From the list below, please choose the three most important service options when choosing your next laptop.

Fast reliable repair service   —
Service at my location   —
Maintenance by the manufacturer 1
Knowledgeable technicians   3
Notification of upgrades   —

same"). Thus, two-way questions may become multiple-choice or rating questions, and these additional responses should be reflected in your revised analysis plan. Dichotomous questions generate nominal data.

Multiple-Choice Question

Multiple-choice questions are appropriate where there are more than two alternatives or where we seek gradations of preference, interest, or agreement; the latter situation also calls for rating questions. While such questions offer more than one alternative answer, they request that the participant make a single choice. Multiple-choice questions can be efficient, but they also present unique design and analysis problems.

Organizations use questionnaires to measure all sorts of activities and attitudes. Kraft used a survey to measure whether its *Food and Family* magazine readers wanted stickers to mark favorite recipe pages. Grocery chain The Kroger Company and Applebee's restaurants use surveys to evaluate customer satisfaction. As reflected in this chart, NOP World wanted to track activity in mobile phone ownership among U.S. teenagers.

Percentage of U.S. Teens Owning Mobile Phones
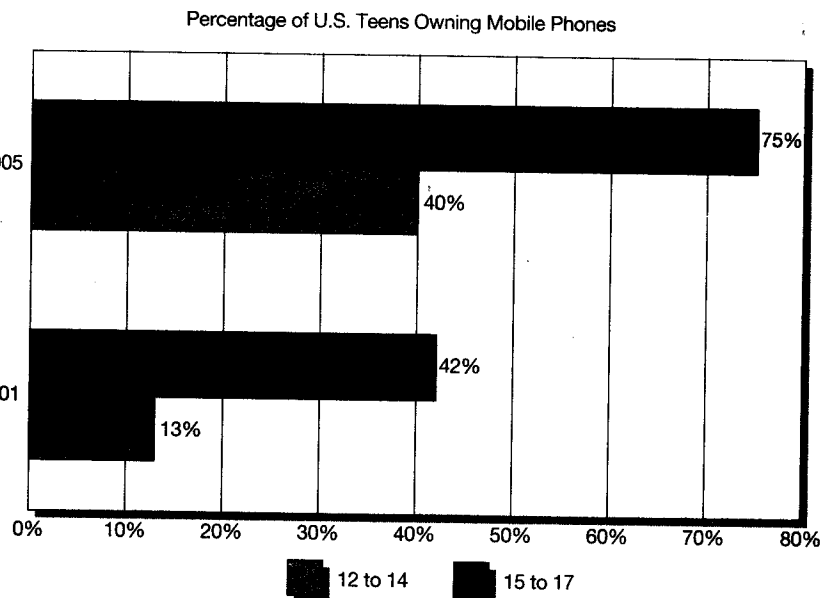
One type of problem occurs when one or more responses have not been anticipated. Assume we ask whether retail mall security and safety rules should be determined by the (1) store managers, (2) sales associates who work at the mall, (3) federal government, or (4) state government. The union has not been mentioned in the alternatives. Many participants might combine this alternative with "sales associates," but others will view "unions" as a distinct alternative. Exploration prior to drafting the measurement question attempts to identify the most likely choices.

A second problem occurs when the list of choices is not exhaustive. Participants may want to give an answer that is not offered as an alternative. This may occur when the desired response is one that combines two or more of the listed individual alternatives. Many participants may believe the store management *and* the sales associates acting jointly should set store safety rules, but the question does not include this response. When the researcher tries to provide for all possible options, choosing from the list of alternatives can become exhausting. We guard against this by discovering the major choices through exploration and pretesting (discussed in detail in Appendix 14b). We may also add the category "other (please specify)" as a safeguard to provide the participant with an acceptable alternative for all other options. In our analysis of responses to a pretested, self-administered questionnaire we may create a combination alternative.

Yet another problem occurs when the participant divides the question of store safety into several questions, each with different alternatives. Some participants may believe rules dealing with air quality in stores should be set by a federal agency while those dealing with aisle obstructions or displays should be set by store management and union representatives. Still others may want store management in conjunction with a sales associate committee to make rules. To address this problem, the instrument designer would need to divide the question. Pretesting should reveal if a multiple-choice question is really a **double-barreled question.**

Another challenge in alternative selection occurs when the choices are not mutually exclusive (the participant thinks two or more responses overlap). In a multiple-choice question that asks students, "Which one of the following factors was most influential in your decision to attend Metro U?" these response alternatives might be listed:

1. Good academic reputation.

2. Specific program of study desired.

3. Enjoyable campus life.

4. Many friends from home attend.

5. High quality of the faculty.

6. Opportunity to play collegiate-level sports.

Some participants might view items 1 and 5 as overlapping, and some may see items 3 and 4 in the same way.

It is also important to seek a fair balance in choices when a participant's position on an issue is unknown. One study showed that an off-balance presentation of alternatives biases the results in favor of the more heavily offered side.[8] If four gradations of alternatives are on one side of an issue and two are offered reflecting the other side, responses will tend to be biased toward the better-represented side. However, researchers may have a valid reason for using an unbalanced array of alternatives. They may be trying to determine the degree of positive (or negative) response, already knowing which side of an issue most participants will choose based on the selection criteria for participation.

It is necessary in multiple-choice questions to present reasonable alternatives—particularly when the choices are numbers or identifications. If we ask, "Which of the following numbers is closest to the number of students enrolled in American colleges and universities today?" these choices might be presented:

1. 75,000

2. 750,000

3. 7,500,000

4. 25,000,000

5. 75,000,000

It should be obvious to most participants that at least three of these choices are not reasonable, given general knowledge about the population of the United States and about the colleges and universities in their hometowns. (The estimated 2003 U.S. population is 290.8 million based on the 2000 census of 281.4 million. The Ohio State University has more than 50,000 students.)

The order in which choices are given can also be a problem. Numeric alternatives are normally presented in order of magnitude. This practice introduces a bias. The participant assumes that if there is a list of five numbers, the correct answer will lie somewhere in the middle of the group. Researchers are assumed to add a couple of incorrect numbers on each side of the correct one. To counteract this tendency to choose the central position, put the correct number at an extreme position more often when you design a multiple-choice question.

Order bias with nonnumeric response categories often leads the participant to choose the first alternative (**primacy effect**) or the last alternative (**recency effect**) over the middle ones. Primacy effect dominates in visual surveys—self-administered via Web or mail—while recency effect dominates in oral surveys—phone and personal interview surveys.[9] Using the *split-ballot technique* can counteract this bias: different segments of the sample are presented alternatives in different orders. To implement this strategy in face-to-face interviews, the researcher would list the alternatives on a card to be handed to the participant when the question is asked. Cards with different choice orders can be alternated to ensure positional balance. The researcher would leave the choices unnumbered on the card so that the participant replies by giving the response category itself rather than its identifying number. It is a good practice to use cards like this any time there are four or more choice alternatives. This saves the interviewer reading time and ensures a more valid answer by keeping the full range of choices in front of the participant. With computer-assisted surveying, the software can be programmed to rotate the order of the alternatives so that each participant receives the alternatives in randomized order (for nonordered scales) or in reverse order (for ordered scales).

In most multiple-choice questions, there is also a problem of ensuring that the choices represent a one-dimensional scale—that is, the alternatives to a given question should represent different aspects of the same conceptual dimension. In the college selection example, the list included features associated with a college that might be attractive to a student. This list, while not exhaustive, illustrated aspects of the concept "college attractiveness factors within the control of the college." The list did not mention other factors that might affect a school attendance decision. Parents and peer advice, local alumni efforts, and one's high school adviser may influence the decision, but these represent a different conceptual dimension of "college attractiveness factors"—those not within the control of the college.

Multiple-choice questions usually generate nominal data. When the choices are numeric alternatives, this response structure may produce at least interval and sometimes ratio data. When the choices represent ordered but unequal, numerical ranges (for example, a question on family income: <$20,000; $20,000–$100,000; >$100,000) or a verbal rating scale (for example, a question on how you prefer your steak prepared: well done, medium well, medium rare, or rare), the multiple-choice question generates ordinal data.

Checklist

When you want a participant to give multiple responses to a single question, you will ask the question in one of three ways: the checklist, rating, or ranking strategy. If relative order is not important, the **checklist** is the logical choice. Questions like "Which of the following factors encouraged you to apply to Metro U? (Check all that apply)" force the participant to exercise a dichotomous response (yes, encouraged; no, didn't encourage) to each factor presented. Of course, you could have asked for the same information with a series of dichotomous selection questions, one for each individual factor, but this would have been both time- and space-consuming. Checklists are more efficient. Checklists generate nominal data.

Rating Question

Rating questions ask the participant to position each factor on a companion scale, either verbal, numeric, or graphic. "Each of the following factors has been shown to have some influence on a student's choice to apply to Metro U. Using your own experience, for each factor please tell us whether the factor was 'strongly influential,' 'somewhat influential,' or 'not at all influential.'" Generally, rating-scale structures generate ordinal data; some carefully crafted scales generate interval data.

It is important to remember that the researcher should represent only one response dimension in rating-scale response options. Otherwise, effectively, you present the participant with a double-barreled question with insufficient choices to reply to both aspects.

Example A: How likely are you to enroll at Metro University?

            (Responses with more than one dimension, ordinal scale)

            (a) extremely likely to enroll

            (b) somewhat likely to enroll

            (c) not likely to apply

            (d) will not apply

Example B: How likely are you to enroll at Metro University?

            (Responses within one dimension, interval scale)

            (a) extremely likely to enroll

            (b) somewhat likely to enroll

            (c) neither likely nor unlikely to enroll

            (d) somewhat unlikely to enroll

            (e) extremely unlikely to enroll

When futurist Arnold Mitchell created the VALS survey for psychographic profiling in 1978, he likely didn't realize that it would become one of the marketers' best-known surveys. The original survey was created "to explain changing U.S. values and lifestyles in the 1970s." Today, VALS is used by marketers to increase the effectiveness of direct marketing efforts, to refine retail merchandise mixes, and to tailor programming for media. It has helped successfully launch new products (by helping select the right people for focus groups, ideation sessions, and concept tests) and been instrumental in brand differentiation strategies and customer retention efforts. In promotion, VALS helps focus brand personality and create ads that use the right language and images that "resonate with each segment."

VALS uses a multi-item survey where each item measures the participant's agreement on a 4-point scale of "mostly disagree," "somewhat disagree," "somewhat agree," and "mostly agree." Participants are grouped into eight major segments based on attitudes and lifestyles. These segments range from *Innovators*—take-charge types who are into status—to *Survivors*—those seeking comfort and security in their limited-focus lives.

You can learn more about the types of projects that use the VALS survey at www.sric-bi.com/VALS/applications.shtml. You may also take your own VALS survey and assess where you fit in the VALS eight-category segmentation scheme at www.sric-bi/vals/survey.shtml.

www.sric-bi.com

Ranking Question

When relative order of the alternatives is important, the **ranking question** is ideal. "Please rank-order your top three factors from the following list based on its influence in encouraging you to apply to Metro U. Use 1 to indicate the most encouraging factor, 2 the next most encouraging factor, etc." The checklist strategy would provide the three factors of influence, but we would have no way of knowing the importance the participant places on each factor. Even in a personal interview, the order in which the factors are mentioned is not a guarantee of influence. Ranking as a response strategy solves this problem.

One concern surfaces with ranking activities. How many presented factors should be ranked? If you listed the 15 brands of potato chips sold in a given market, would you have the participant rank all 15 in order of preference? In most instances it is helpful to remind yourself that while participants may have been selected for a given study due to their experience or likelihood of having desired information, this does not mean that they have knowledge of all conceivable aspects of an issue, but only of some. It is always better to have participants rank only those elements with which they are familiar. For this reason, ranking questions might appropriately follow a checklist question that identifies the objects of familiarity. If you want motivation to remain strong, avoid asking a participant to rank more than seven items even if your list is longer. Ranking generates ordinal data.

All types of response strategies have their advantages and disadvantages. Several different strategies are often found in the same questionnaire, and the situational factors mentioned earlier are the major guides in this matter. There is a tendency, however, to use closed questions instead of the more flexible open-ended type. Exhibit 14-7 summarizes some important considerations in choosing between the various response strategies.

Sources of Existing Questions

The tools of data collection should be adapted to the problem, not the reverse. Thus, the focus of this chapter has been on crafting an instrument to answer specific investigative questions. But inventing, refining, and pretesting questions demands considerable time and effort. For some topics, a careful review of the related

> **Exhibit 14-7** Characteristics of Response Strategies

Characteristic	Dichotomous	Multiple Choice	Checklist	Rating	Ranking	Free Response
Type of Scale	Nominal	Nominal, ordinal, or ratio	Nominal	Ordinal or interval	Ordinal	Nominal or ratio
Desired Number of Participant Answers	1	1	10 or fewer	1 per item	7 or fewer	1

literature and an examination of existing instrument sourcebooks can shorten this process. Increasingly, companies that specialize in survey research maintain a question bank of pretested questions. In the opening vignette, Sally was accessing Jason's question bank.

A review of literature will reveal instruments used in similar studies that may be obtained by writing to the researchers or, if copyrighted, may be purchased through a clearinghouse. Instruments also are available through compilations and sourcebooks. While these tend to be oriented to social science applications, they are a rich source of ideas for tailoring questions to meet a manager's needs. Several compilations are recommended; we have suggested them in Exhibit 14-8.[10]

Borrowing items from existing sources is not without risk. It is quite difficult to generalize the reliability and validity of selected questionnaire items or portions of a questionnaire that have been taken out of the original context. Researchers whose questions or instruments you borrow may not have reported sampling and testing procedures needed to judge the quality of the measurement scale. Just because Jason has a satisfaction scale in the question bank used for the CardioQuest survey does not mean the question will be appropriate for the Albany Outpatient Laser Clinic. Sally would need to know the intended purpose of the CardioQuest study and the time of construction, as well as the results of pretesting, to determine the reliability and validity of its use in the Albany study. Even then she would be wise to pretest the question in the context of her Albany survey.

Language, phrasing, and idioms can also pose problems. Questions tend to age or become outdated and may not appear (or sound) as relevant to the participant as freshly worded questions. Integrating previously used and customized questions is problematic. Often adjacent questions in one questionnaire are relied on to carry context. If you select one question from a contextual series, the borrowed question is left without its necessary meaning.[11] Whether an instrument is constructed with designed questions or adapted with questions borrowed or licensed from others, pretesting is expected.

> Drafting and Refining the Instrument

This section reflects Phase 3 in Exhibit 14-1.　As depicted in Exhibit 14-9, Phase 3 of instrument design—drafting and refinement—is a multistep process:

1. Develop the participant-screening process (done especially with personal or phone surveys, but also with early notification procedures with e-mail and Web surveys), along with the introduction.

> **Exhibit 14-8** Sources of Questions

Printed Sources		
Author(s)	**Title**	**Source**
William Bearden and R. Netemeyer	*Handbook of Marketing Scales: Multi-Item Measures for Marketing and Consumer Behavior Research*	London: Sage Publications, Inc., 2001
John P. Robinson, Philip R. Shaver, and Lawrence S. Wrightsman	*Measures of Personality and Social-Psychological Attitudes*	San Diego, CA: Academic Press, 1990, 1999
George H. Gallup Jr., ed.	*The Gallup Poll: Public Opinion 1998*	Wilmington, DE: Scholarly Resources, 1998
Elizabeth H. Hastings and Philip K. Hastings, eds.	*Index to International Public Opinion, 1986–1987*	Westport, CT: Greenwood Publishing Group, September 1988
Philip E. Converse, Jean D. Dotson, Wendy J. Hoag, and William H. McGee III, eds.	*American Social Attitudes Data Sourcebook, 1947–1978*	Cambridge, MA: Harvard University Press, 1980
National Opinion Research Center	*General Social Surveys 1972–2000: Cumulative Code Book*	Ann Arbor, MI: ICPSR, 2000
Web Sources		
Interuniversity Consortium for Political and Social Research (general social survey)	www.icpsr.umich.edu:8080/GSS/index.html	
Online Survey Research/Public Opinion Centers (a world listing)	www.ku.edu/cwis/units/coms2/po/	
The Odum Institute (houses the Louis Harris Opinion Polls)	www.irss.unc.edu	

> **Exhibit 14-9** Flowchart for Instrument Design: Phase 3
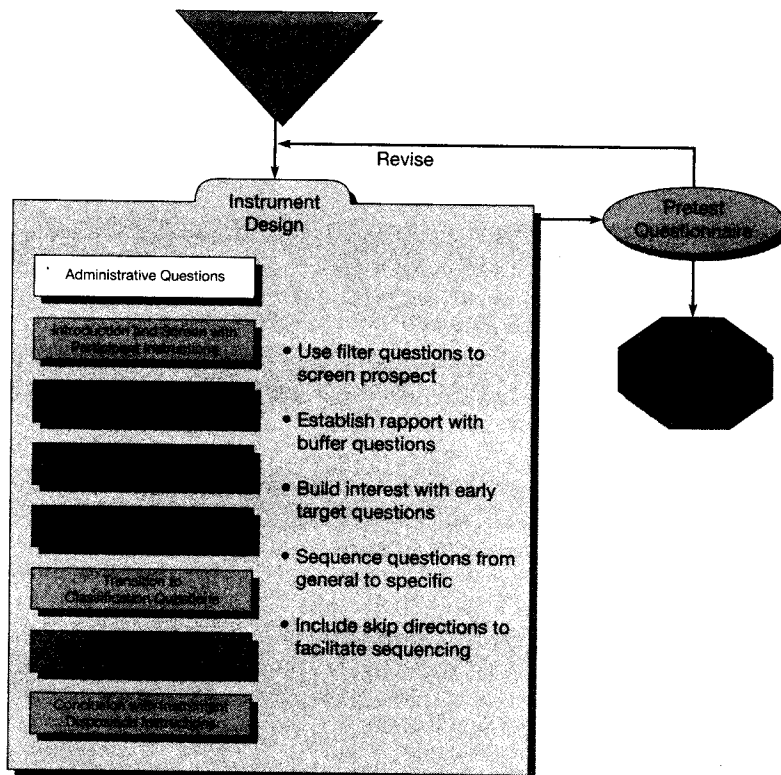
2. Arrange the measurement question sequence:

  a. Identify groups of target questions by topic.

  b. Establish a logical sequence for the question groups and questions within groups.

  c. Develop transitions between these question groups.

3. Prepare and insert instructions—for the interviewer or participant—including termination instructions, skip directions, and probes.

4. Create and insert a conclusion, including a survey disposition statement.

5. Pretest specific questions and the instrument as a whole.

Participant Screening and Introduction

The introduction must supply the sample unit with the motivation to participate in the study. It must reveal enough about the forthcoming questions, usually by revealing some or all of the topics to be covered, for participants to judge their interest level and their ability to provide the desired information. In any communication study, the introduction also reveals the amount of time participation is likely to take. The introduction also reveals the research organization or sponsor (unless the study is disguised) and possibly the objective of the study. In personal or phone interviews as well as in e-mail and Web surveys, the introduction usually contains one or more **screen questions** or filter questions to determine if the potential participant has the knowledge or experience necessary to participate in the study. At a minimum, a phone or personal interviewer will introduce himself or herself to help establish critical rapport with the potential participant. Exhibit 14-10 provides a sample introduction and other components of a telephone study of nonparticipants to a self-administered mail survey.

Maximum Online Survey Length Prior to Abandonment

More than 20 minutes,
13.3%

16–20 minutes, 9.0%

5 minutes or less,
33.9%

11–15 minutes, 15.7%
15 minutes or less, 77.6%

6–10 minutes, 28%
10 minutes or less, 61.9%
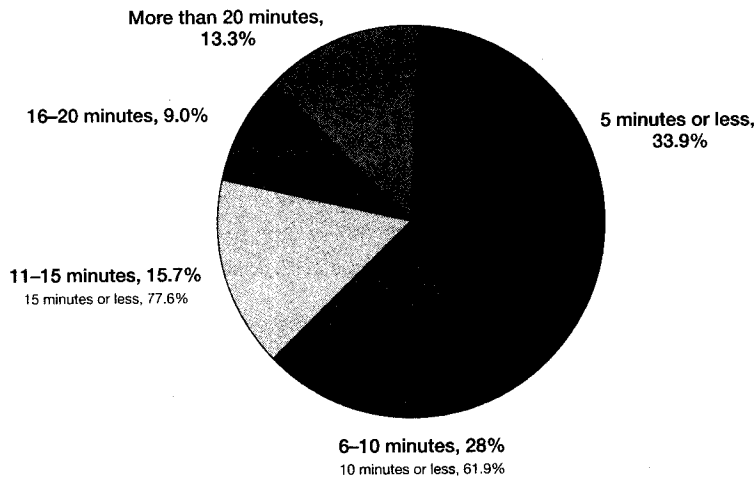
>picprofile

As marketing resistance rises and survey cooperation declines, survey length is of increasing concern. InsightExpress studied the Web survey process and revealed that people taking Web surveys prefer shorter to longer surveys, consistent with what we know about phone and intercept survey participants. While 77 percent were likely to complete a survey that took 15 minutes or less, almost one in three participants needed a survey to be 5 minutes or less for full completion. As participating in online surveys loses its novelty, prospective participants are likely to become even more reluctant to give significant time to the survey process. Therefore, it is critical that researchers ask only what is necessary. www.insightexpress.com

>snapshot

Does Direct-Response TV Influence Purchase Behavior?

The year 2004 marked the 20th anniversary of the infomercial, so the Electronic Retailing Association (ERA) decided the time was right for a tracking study. Three types of direct-response television (DRTV) tempt viewers to purchase products directly: the direct-response commercial (30- to 120-second TV commercial), the infomercial (30-minute paid program), and live shopping shows (like QVC, Home Shopping Network, Shop at Home, and ShopNBC). According to a 300-participant telephone survey conducted by Leisure Trends Group for ERA, since 1996 the percentage of those 18 or older viewing the three types of DRTV has declined 11 percent but now reaches 63 percent. This translates to 136.2 million viewers annually. Among 16- to 24-year-olds, viewership has increased compared to that in the previous year. ERA sees this as good news given that television viewing overall by persons in this group is declining as they turn increasingly to the Internet for entertainment. Overall, 37 percent of Americans say they have been enticed to buy a product they saw advertised on TV—39 percent of infomercial viewers, 34 percent of live shopping show viewers, and 32 percent of direct-response commercial viewers. These percentages are all lower than they were in a comparative study in 1994. Of those who purchased, 76 percent purchased from a store or the Internet, while 59 percent dialed the phone number given on the TV. If you were considering DRTV as a distribution option, would this study convince you to use DRTV?

www.retailing.org; www.leisuretrends.com

Measurement Question Sequencing

The design of survey questions is influenced by the need to relate each question to the others in the instrument. Often the content of one question (called a **branched question**) assumes other questions have been

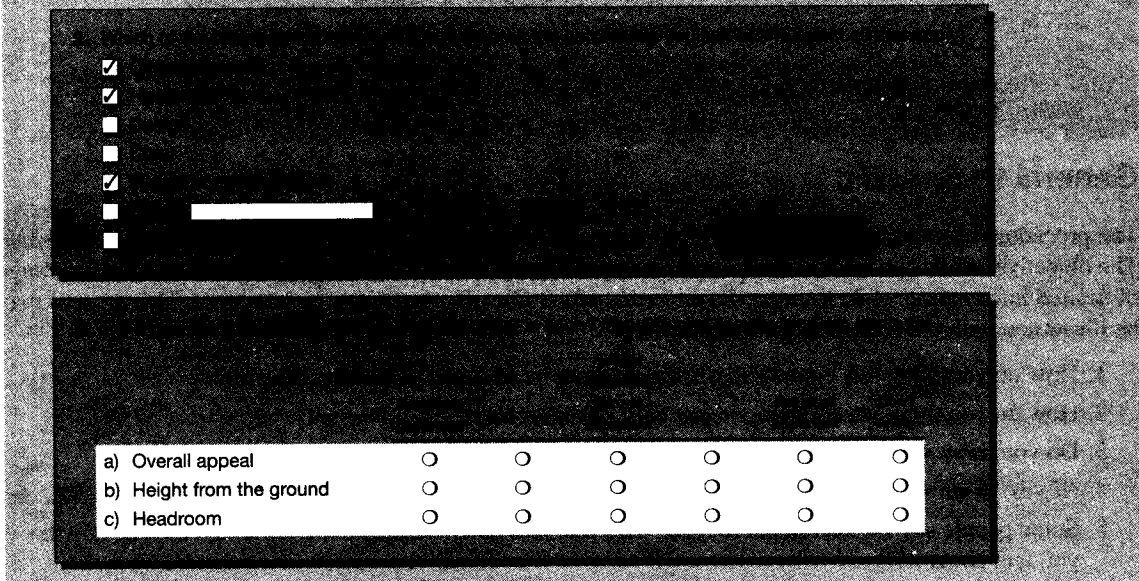**> Exhibit 14-10** Sample Components of Communication Instruments

Component	Example
Introduction	Good evening. May I please speak with (name of participant)? Mr. (participant's last name), I'm (your name), calling on behalf of MindWriter Corporation. You recently had your MindWriter laptop serviced at our CompleteCare Center. Could you take five minutes to tell us what you thought of the service provided by the center?
Instructions for . . . a. Terminating (following filter or screen question)	I'm sorry, today we are only talking with individuals who eat cereal at least three days per week, but thank you for speaking with me. (Pause for participant reply) Good-bye.
c. Skip directions (between questions or groups of questions)	3. Did you purchase boxed cereal in the last 7 days? o Yes o No (skip to question 7)
Conclusion a. Phone or personal interview	That's my last question. Your insights and the ideas of other valuable customers will help us to make the CompleteCare program the best it can be.
b. Self-administered (usually precedes the disposition instructions)	Thank you for talking with us this evening. (Pause for participant reply) Good evening. Thank you for sharing your ideas about the CompleteCare program. Your insights will help us serve you better.

asked and answered. The psychological order of the questions is also important; question sequence can encourage or discourage commitment and promote or hinder the development of researcher-participant rapport.

The basic principle used to guide sequence decisions is this: The nature and needs of the participant must determine the sequence of questions and the organization of the interview schedule. Four guidelines are suggested to implement this principle:

1. The question process must quickly awaken interest and motivate the participant to participate in the interview. Put the more interesting topical target questions early. Leave classification questions (e.g., age, family size, income) not used as filters or screens to the end of the survey.

2. The participant should not be confronted by early requests for information that might be considered personal or ego-threatening. Put questions that might influence the participant to discontinue or terminate the questioning process near the end.

3. The questioning process should begin with simple items and then move to the more complex, as well as move from general items to the more specific. Put taxing and challenging questions later in the questioning process.

4. Changes in the frame of reference should be small and should be clearly pointed out. Use transition statements between different topics of the target question set.

One of the attractions of using a Web survey is the ease with which participants follow branching questions immediately customized to their response patterns. In this survey, participants were shown several pictures of a prototype vehicle. Those who responded to question 2 by selecting one or more of the attributes in the checklist question were sequenced to a version of qu... question 2. Note also that in question 3 the researcher chose ... indicate ... or she had no opinion ("Don't know") on the issue of level of importance.

a) Overall appeal	O	O	O	O	O	O
b) Height from the ground	O	O	O	O	O	O
c) Headroom	O	O	O	O	O	O

Awaken Interest and Motivation

We awaken interest and stimulate motivation to participate by choosing or designing questions that are attention-getting and not controversial. If the questions have human-interest value, so much the better. It is possible that the early questions will contribute valuable data to the major study objective, but their major task is to overcome the motivational barrier.

Sensitive and Ego-Involving Information

Regarding the introduction of sensitive information too early in the process, two forms of this error are common. Most studies need to ask for personal classification information about participants. Participants normally will provide these data, but the request should be made at the end of the survey. If made at the start of the survey, it often causes participants to feel threatened, dampening their interest and motivation to continue. It is also dangerous to ask any question at the start that is too personal. For example, participants in one survey were asked whether they suffered from insomnia. When the question was asked immediately after the interviewer's introductory remarks, about 12 percent of those interviewed admitted to having insomnia. When a matched sample was asked the same question after two **buffer questions** (neutral questions designed chiefly to establish rapport with the participant), 23 percent admitted suffering from insomnia.[12]

Simple to Complex

Deferring complex questions or simple questions that require much thought can help reduce the number of "don't know" responses that are so prevalent early in interviews.

> **Exhibit 14-11** Question Sequencing

Question	Percent Answering Yes	
	A. Asked First	B. Asked First
A. Should the United States permit its citizens to join the French and British armies?	45%	40%
B. Should the United States permit its citizens to join the German army?	31	22

General to Specific

The procedure of moving from general to more specific questions is sometimes called the *funnel approach*. The objectives of this procedure are to learn the participant's frame of reference and to extract the full range of desired information while limiting the distortion effect of earlier questions on later ones. This process may be illustrated with the following series of questions:

1. How do you think this country is getting along in its relations with other countries?
2. How do you think we are doing in our relations with Iran?
3. Do you think we ought to be dealing with Iran differently than we are now?
4. (If yes) What should we be doing differently?
5. Some people say we should get tougher with Iran and others think we are too tough as it is; how do you feel about it?[13]

< **You might find it valuable to refer to Exhibit 8-6, "The Interview Question Hierarchy," page 207.**

The first question introduces the general subject and provides some insight into the participant's frame of reference. The second question narrows the concern to a single country, while the third and fourth seek views on how the United States should deal with Iran. The fifth question illustrates a specific opinion area and would be asked only if this point of toughness had not been covered in earlier responses. Question 4 is an example of a branched question; the response to the previous question determines whether or not question 4 is asked of the participant.

There is also a risk of interaction whenever two or more questions are related. Question-order influence is especially problematic with self-administered questionnaires, because the participant is at liberty to refer back to questions previously answered. In an attempt to "correctly align" two responses, accurate opinions and attitudes may be sacrificed. Computer-administered and Web surveys have largely eliminated this problem.

The two questions shown in Exhibit 14-11 were asked in a national survey at the start of World War II.[14] Apparently, some participants who first endorsed enlistment with the Allies felt obliged to extend this privilege to joining the German army. Where the decision was first made against joining the German army, a percentage of the participants felt constrained from approving the option to join the Allies.

Question Groups and Transitions

The last question-sequencing guideline suggests arranging questions to minimize shifting in subject matter and frame of reference. Participants often interpret questions in the light of earlier questions and miss shifts of perspective or subject unless they are clearly stated. Participants fail to listen carefully and frequently jump to conclusions about the import of a given question before it is completely stated. Their answers are strongly influenced by their frame of reference. Any change in subject by the interviewer may not register with them unless it is made strong and obvious. Most questionnaires that cover a range of topics are divided into

sections with clearly defined transitions between sections to alert the participant to the change in frame of reference. Exhibit 14-12 provides a sample of a transition in the MindWriter CompleteCare study when measurement questions changed from service-related questions to personal and family-related questions.

Instructions

Instructions to the interviewer or participant attempt to ensure that all participants are treated equally, thus avoiding building error into the results. Two principles form the foundation for good instructions: clarity and courtesy. Instruction language needs to be unfailingly simple and polite.

Instruction topics include those for:

- *Terminating an unqualified participant*—defining for the interviewer how to terminate an interview when the participant does not correctly answer the screen or filter questions.

- *Terminating a discontinued interview*—defining for the interviewer how to conclude an interview when the participant decides to discontinue.

- *Moving between questions on an instrument*—defining for an interviewer or participant how to move between questions or topic sections of an instrument (*skip directions*) when movement is dependent on the specific answer to a question or when branched questions are used.

- *Disposing of a completed questionnaire*—defining for an interviewer or participant completing a self-administered instrument how to submit the completed questionnaire.

In a self-administered questionnaire, instructions must be contained within the survey instrument. Personal interviewer instructions sometimes are in a document separate from the questionnaire (a document thoroughly discussed during interviewer training) or are distinctly and clearly marked (highlighted, printed in colored ink, or boxed on the computer screen or in a pop-up window) on the data collection instrument itself. Sample instructions are presented in Exhibit 14-12.

Conclusion

The role of the conclusion is to leave the participant with the impression that his or her involvement has been valuable. Subsequent researchers may need this individual to participate in new studies. If every interviewer or instrument expresses appreciation for participation, cooperation in subsequent studies is more likely. A sample conclusion is shown in Exhibit 14-12.

Overcoming Instrument Problems

There is no substitute for a thorough understanding of question wording, question content, and question sequencing issues. However, the researcher can do several things to help improve survey results, among them:

- Build rapport with the participant.
- Redesign the questioning process.
- Explore alternative response strategies.
- Use methods other than surveying to secure the data.
- Pretest all the survey elements.

> **See Appendix 14b on pretesting for coverage of these final two bullets.**

Build Rapport with the Participant

Most information can be secured by direct undisguised questioning if rapport has been developed. Rapport is particularly useful in building participant interest in the project, and the more interest participants have, the more cooperation they will give. One can also overcome participant unwillingness by providing some material compensation for cooperation. This approach has been especially successful in mail surveys and is increasingly used in Web surveys.

The assurance of confidentiality also can increase participants' motivation. One approach is to give discrete assurances, both by question wording and by interviewer comments and actions, that all types of behavior, attitudes, and positions on controversial or sensitive subjects are acceptable and normal. Where you can say so truthfully, guarantee that participants' answers will be used only in combined statistical totals (aggregate data), not matched to an individual participant. If participants are convinced that their replies contribute to some important purpose, they are more likely to be candid, even about taboo topics. If a researcher's organization uses an Institutional Review Board to review surveys before use, the board may require an instruction indicating that any response—in fact, participation—is voluntary. This is especially important where surveys are used with internal publics (employees).

Redesign the Questioning Process

You can redesign the questioning process to improve the quality of answers by modifying the administrative process and the response strategy. We might show that confidentiality is indispensable to the administration of the survey by using a group administration of questionnaires, accompanied by a ballot-box collection procedure. Even in face-to-face interviews, the participant may fill in the part of the questionnaire containing sensitive information and then seal the entire instrument in an envelope. While this does not guarantee confidentiality, it does suggest it.

We can also develop appropriate questioning sequences that will gradually lead a participant from "safe" questions to those that are more sensitive. As already noted in our discussion of disguised questions, indirect questioning (using projective techniques) is a widely used approach for securing opinions on sensitive topics. The participants are asked how "other people" or "people around here" feel about a topic. It is assumed the participants will reply in terms of their own attitudes and experiences, but this outcome is hardly certain. Indirect questioning may give a good measure of the majority opinion on a topic but fail to reflect the views either of the participant or of minority segments.

With certain topics, it is possible to secure answers by using a proxy code. When we seek family income groups, we can hand the participant a card with income brackets like these:

A. Under $25,000 per year.

B. $25,000 to $49,999 per year.

C. $50,000 to $74,999 per year.

D. $75,000 and over per year.

The participant is then asked to report the appropriate bracket as either A, B, C, or D. For some reason, participants are more willing to provide such an obvious proxy measure than to verbalize actual dollar values.

Explore Alternative Response Strategies

When drafting the original question, try developing positive, negative, and neutral versions of each type of question. This practice dramatizes the problems of bias, helping you to select question wording that minimizes such problems. Sometimes use an extreme version of a question rather than the expected one.

>closeup
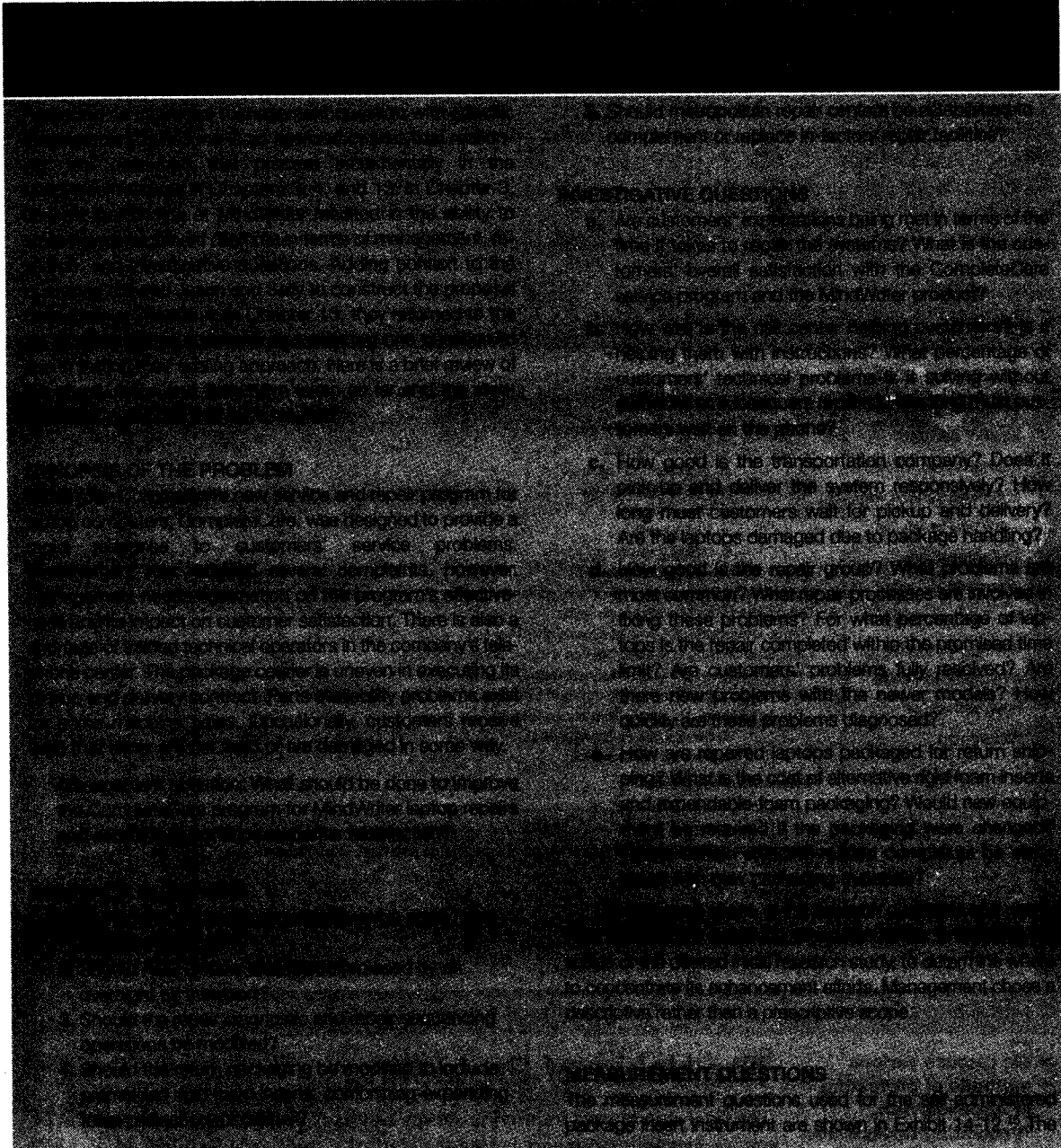
Minimize nonresponses to particular questions by recognizing the sensitivity of certain topics. In a self-administered instrument, for example, asking a multiple-choice question about income or age, where incomes and ages are offered in ranges, is usually more successful than using a free-response question (such as "What is your age, please? _____ ").

< We discussed the use of similar unobtrusive measures in Chapter 9.

>cont'd

[illegible obscured text block]

MindWriter personal	Met few expectations	Met some expectations	Met most expectations	Met all expectations	Exceeded expectations
computers offer you	1	2	3	4	5

MindWriter personal computers offer you ease of use and maintenance. When you need service, we want you to rely on **CompleteCare,** wherever you may be. That's why we're asking you to take a moment to tell us how well we've served you.						
1. Telephone assistance with your problem:						
	a. Responsiveness	1	2	3	4	5
	b. Technical competence	1	2	3	4	5
2. The courier service's effectiveness:						
	a. Arrangements	1	2.	3	4	5
	b. Pickup speed	1	2	3	4	5
	c. Delivery speed	1	2	3	4	5
3. Speed of the overall repair process.		1	2	3	4	5
4. Resolution of the problem that prompted service/repair.		1	2	3	4	5
5. Condition of your MindWriter on arrival.		1	2	3	4	5
6. Overall impression of CompleteCare's effectiveness.		1	2	3	4	5
7. Likelihood of using CompleteCare on another occasion.						
	(1 = very unlikely 3 = neither likely nor unlikely 5 = very likely)	1	2	3	4	5
8. Likelihood of repurchasing a MindWriter based on:						
	(1 = very unlikely 3 = neither likely nor unlikely 5 = very likely)					
	a. Service/repair experience	1	2	3	4	5
	b. Product performance	1	2	3	4	5

Comments/Suggestions: _____

_____

_____

How may we contact you to follow up on any problems you have experienced?

_____ (_____) _____
Last Name   First Name   Phone

_____
City   State   Zip

> We discuss various methods of pretesting in Appendix 14b, "Pretesting Options and Discoveries."

The Value of Pretesting

The final step toward improving survey results is **pretesting,** the assessment of questions and instruments before the start of a study (see Exhibits 14-1, 14-2, and 14-9). There are abundant reasons for pretesting individual questions, questionnaires, and interview schedules: (1) discovering ways to increase participant interest, (2) increasing the likelihood that participants will remain engaged to the completion of the survey, (3) discovering question content, wording, and sequencing problems, (4) discovering target question groups where researcher training is needed, and (5) exploring ways to improve the overall quality of survey data.

Most of what we know about pretesting is prescriptive. According to contemporary authors, there are no general principles of good pretesting, no systematization of practice, no consensus about expectations, and we rarely leave records for each other. How a pretest was conducted, what investigators learned from it, how they redesigned their questionnaire on the basis of it—these matters are reported only sketchily in research reports, if at all.[16]

Nevertheless, pretesting not only is an established practice for discovering errors but also is useful for training the research team. Ironically, professionals who have participated in scores of studies are more likely to pretest an instrument than is a beginning researcher hurrying to complete a project. Revising questions five or more times is not unusual. Yet inexperienced researchers often underestimate the need to follow the design-test-revise process.

>summary

1 The instrument design process starts with a comprehensive list of investigative questions drawn from the management-research question hierarchy. Instrument design is a three-phase process with numerous issues within each phase: (a) developing the instrument design strategy, (b) constructing and refining the measurement questions, and (c) drafting and refining the instrument.

2 Several choices must be made in designing a communication study instrument. Surveying can be a face-to-face interview, or it can be much less personal, using indirect media and self-administered questionnaires. The questioning process can be unstructured, as in an IDI, or the questions can be clearly structured. Responses may be unstructured and open-ended or structured with the participant choosing from a list of possibilities. The degree to which the objectives and intent of the questions should be disguised must also be decided.

3 Instruments obtain three general classes of information. Target questions address the investigative questions and are the most important. Classification questions concern participant characteristics and allow participants' answers to be grouped for analysis. Administrative questions identify the participant, interviewer, and interview location and conditions.

4 Question construction involves three critical decision areas. They are (a) question content, (b) question wording, and (c) response strategy. Question content should pass the following tests: Should the question be asked? Is it of proper scope? Can and will the participant answer adequately?

Question wording difficulties exceed most other sources of distortion in surveys. Retention of a question should be confirmed by answering these questions: Is the question stated in terms of a shared vocabulary? Does the vocabulary have a single meaning? Does the question contain misleading assumptions? Is the wording biased? Is it correctly personalized? Are adequate alternatives presented?

The study's objective and participant factors affect the decision of whether to use open-ended or closed questions. Each response strategy generates a specific level of data, with available statistical procedures for each scale type influencing the desired response strategy. Participant factors include level of information about the topic, degree to which the topic has been thought through, ease of communication, and motivation to share information. The decision is also affected by the interviewer's perception of participant factors.

Both dichotomous response and multiple-choice questions are valuable, but on balance the latter are preferred if only because few questions have only two possible answers. Checklist, rating, and ranking strategies are also common.

5 Question sequence can drastically affect participant willingness to cooperate and the quality of responses. Generally, the sequence should begin with efforts to awaken the participant's interest in continuing the interview. Early questions should be simple rather than complex, easy rather than difficult, nonthreatening, and obviously germane to the announced objective of the study. Frame-of-reference changes should be minimal, and questions should be sequenced so that early questions do not distort replies to later ones.

6 Sources of questions for the construction of questionnaires include the literature on related research and sourcebooks of scales and questionnaires. Borrowing items has attendant risks, such as time and situation-specific problems or reliability and validity. Incompatibility of language and idiom also needs to be considered.

>key terms

administrative question 364	dummy table 362	rating question 372
branched question 377	free-response question 367	recency effect 371
buffer question 379	interview schedule 363	screen question (filter question) 376
checklist 372	leading question 367	
classification question 364	multiple-choice question 369	structured response 367
dichotomous question 368	pretesting 384	target question 364
disguised question 360	primacy effect 371	structured, 364
double-barreled question 370	ranking question 373	unstructured, 364
		unstructured response 367

>discussion questions

Terms in Review

1 Distinguish between:
   a Direct and indirect questions.
   b Open-ended and closed questions.
   c Research, investigative, and measurement questions.
   d Alternative response strategies.

2 Why is the survey technique so popular? When is it not appropriate?

3 What special problems do open-ended questions have? How can these be minimized? In what situations are open-ended questions most useful?

4 Why might a researcher wish to disguise the objective of a study?

5 One of the major reasons why survey research may not be effective is that the survey instruments are less useful than they should be. What would you say are the four possible major faults of survey instrument design?

6 Why is it desirable to pretest survey instruments? What information can you secure from such a pretest? How can you find the best wording for a question on a questionnaire?

7 One design problem in the development of survey instruments concerns the sequence of questions. What suggestions would you give to researchers designing their first questionnaire?

8 One of the major problems facing the designer of a survey instrument concerns the assumptions made. What are the major "problem assumptions"?

Making Research Decisions

9 Below are six questions that might be found on questionnaires. Comment on each as to whether or not it is a good question. If it is not, explain why.

(Assume that no lead-in or screening questions are required. Judge each question on its own merits.)
   a Do you read *National Geographic* magazine regularly?
   b What percentage of your time is spent asking for information from others in your organization?
   c When did you first start chewing gum?
   d How much discretionary buying power do you have each year?
   e Why did you decide to attend Big State University?
   f Do you think the president is doing a good job now?

10 In a class project, students developed a brief self-administered questionnaire by which they might quickly evaluate a professor. One student submitted the following instrument. Evaluate the questions asked and the format of the instrument.

Professor Evaluation Form

1. Overall, how would you rate this professor?
   ☐ Good   ☐ Fair   ☐ Poor
2. Does this professor
   a. Have good class delivery? _____
   b. Know the subject? _____
   c. Have a positive attitude toward the subject?
      _____
   d. Grade fairly? _____
   e. Have a sense of humor? _____
   f. Use audiovisuals, case examples, or other classroom aids? _____
   g. Return exams promptly? _____
3. What is the professor's strongest point? _____
4. What is the professor's weakest point? _____
5. What kind of class does the professor teach?
   _____
6. Is this course required? _____

7. Would you take another course from this professor? _____

11 Assume the American Society of Training Directors is studying its membership in order to enhance member benefits and attract new members. Below is a copy of a cover letter and mail questionnaire received by a member of the society. Please evaluate the usefulness and tone of the letter and the questions and format of the instrument.

Dear ASTD Member:

The ASTD is evaluating the perception of value of membership among its members. Enclosed is a short questionnaire and a return envelope. I hope you will take a few minutes and fill out the questionnaire as soon as possible, as the sooner the information is returned to me, the better.

Sincerely,
Director of Membership

Questionnaire

Directions: Please answer as briefly as possible.

1. With what company did you enter the field of training? _____

2. How long have you been in the field of training? _____

3. How long have you been in the training department of the company with which you are presently employed? _____

4. How long has the training department in your company been in existence? _____

5. Is the training department a subset of another department? If so, what department? _____

6. For what functions (other than training) is your department responsible? _____

7. How many people, including yourself, are in the training department of your company (local plant or establishment)? _____

8. What degrees do you hold and from what institutions? _____

   Major _____ Minor _____

9. Why were you chosen for training? What special qualifications prompted your entry into training? _____

10. What experience would you consider necessary for an individual to enter into the field of training with your company? Include both educational requirements and actual experience. _____

Bringing Research to Life

12 Design the introduction of the Albany Outpatient Laser Clinic survey, assuming it will continue to be a self-administered questionnaire.

13 To evaluate whether presurgery patient attitudes affect recovery and ultimate patient satisfaction with the Albany Outpatient Laser Clinic, design a question for the self-administered survey. (You may wish to review the opening vignettes in this chapter and Chapter 10.)

From Concept to Practice

14 Using Exhibits 14-1, 14-4, and 14-9, develop the flowchart for the Albany Outpatient Laser Clinic study in the opening vignette.

>wwwexercises

1 Volunteer to participate in a survey on the Web. You can use a search engine to find such sites (keyword: *survey panel*), or visit **http://www.dreamwater.org/surveys/** and click on the links provided to such panels as Greenfield Online and NFO. What questionnaire design decisions were made for the survey you took?

2 You can experience a sample Web survey by visiting the InsightExpress shared surveys Web page (**http://www. insightexpress.com**). Look for "Topline Reports," and click on "access results now."

cases*

Can Research Rescue the Red Cross?	Inquiring Minds Want to Know—NOW!
	Mastering Teacher Leadership

**NCRCC: Teeing Up and New
Strategic Direction**

USTA: Come Out Swinging

T-Shirt Designs

* All cases appear on the text CD; you will find abstracts of these cases in the Case Abstracts section of this text.
Video cases are indicated with a video icon.

>appendix14a

Crafting Effective Measurement Questions

Numerous issues influence whether the questions we ask on questionnaires generate the decision-making data that managers sorely need. Each of the issues summarized in Exhibit 14-5 is developed more fully here.

Question Content

Should This Question Be Asked?

Purposeful versus Interesting Questions that merely produce "interesting information" cannot be justified on either economic or research grounds. Challenge each question's function. Does it contribute significant information toward answering the research question? Will its omission limit or prevent the thorough analysis of other data? Can we infer the answer from another question? A good question designer knows the value of learning more from fewer questions.

Is the Question of Proper Scope and Coverage?

Incomplete or Unfocused We can test this content issue by asking, "Will this question reveal all we need to know?" We sometimes ask participants to reveal their motivations for particular behaviors or attitudes by asking them, "Why?" This simple question is inadequate to probe the range of most causal relationships. When studying product use behavior, for example, we learn more by directing two or three questions on product use to the heavy-use consumer and only one question to the light user.

Questions are also inadequate if they do not provide the information you need to interpret responses fully. If you ask about the Albany Clinic's image for quality patient care, do different groups of patients or those there for the first versus the third time have different attitudes? To evaluate relative attitudes, do you need to ask the same question about other companies? In the original Albany Clinic survey, participants were asked, "Have you ever had or been treated for a recent cold or flu?" If participants answer yes, what exactly have they told the researcher that would be of use to the eye surgeon? Wouldn't it be likely that the surgeon is interested in med-ication taken to treat colds or flu within, say, the prior 10 days? This question also points to two other problems of scope and coverage: the double-barreled question and the imprecise question.

Double-Barreled Questions Does the question request so much content that it should be broken into two or more questions? While reducing the overall number of questions in a study is highly desirable, don't try to ask double-barreled questions. The Albany Clinic question about flu ("Have you ever had or been treated for a recent cold or flu?") fires more than two barrels. It asks four questions in all (Ever had cold? Ever had flu? Been treated for cold? Been treated for flu?).

Here's another common example posed to menswear retailers: "Are this year's shoe sales and gross profits higher than last year's?" Couldn't sales be higher with stagnant profits, or profits higher with level or lower sales? This second example is more typical of the problem of double-barreled questions.

A less obvious double-barreled question is the question we ask to identify a family's or a group's TV station preference. Since a single station is unlikely, a better question would ask the station preference of each family member separately or, alternatively, screen for the group member who most often controls channel selection on Monday evenings during prime time. Also, it's highly probable that no one station would serve as an individual's preferred station when we cover a wide range of time (8 to 11 p.m.). This reveals another problem, the imprecise question.

Precision To test a question for precision, ask, "Does the question ask precisely what we want and need to know?" We sometimes ask for a participant's income when we really want to know the family's total annual income before taxes in the past calendar year. We ask what a participant purchased "last week" when we really want to know what he or she purchased in a "typical 7-day period during the past 90 days." The Albany Clinic's patients were asked for cold and flu history during the time frame "ever." It is hard to imagine an adult who has never experienced a cold or flu and equally hard to assume an adult hasn't been treated for one or both at some time in his or her life.

A second precision issue deals with common vocabulary between researcher and participant. To test your question for this problem, ask, "Do I need to offer operational definitions of concepts and constructs used in the question?"

Can the Participant Answer Adequately?

Time for Thought Although the question may address the topic, is it asked in such a way that the participant will be able to frame an answer, or is it reasonable to assume that the participant can determine the answer? This is also a question that drives sample design, but once the ideal sample unit is determined, researchers often assume that participants who fit the sample profile have all the answers, preferably on the tips of their tongues. To frame a response to some questions takes time and thought; such questions are best left to self-administered questionnaires.

Participation at the Expense of Accuracy Participants typically want to cooperate in interviews; thus they assume giving any answer is more helpful than denying knowledge of a topic. Their desire to impress the interviewer may encourage them to give answers based on no information. A classic illustration of this problem occurred with the following question:[1] "Which of the following statements most closely coincides with your opinion of the Metallic Metals Act?" The response pattern shows that 70 percent of those interviewed had a fairly clear opinion of the Metallic Metals Act; however, there is no such act. The participants apparently assumed that if a question was asked, they should provide an answer. Given reasonablesounding choices, they selected one even though they knew nothing about the topic.

To counteract this tendency to respond at any cost, *filter* or *screen questions* are used to qualify a participant's knowledge. If the MindWriter service questionnaire is distributed via mail to all recent purchasers of MindWriter products, we might ask, "Have you required service for your laptop since its purchase?" Only those for whom service was provided could supply the detail and scope of the responses indicated in the investigative question list. If such a question is asked in a phone interview, we would call the question a *screen,* because it is being used to determine whether the person on the other end of the phone line is a qualified sample unit. This same question asked on a computer-administered questionnaire would likely *branch* or *skip* the participant to a series of classification questions.

Assuming that participants have prior knowledge or understanding may be risky. The risk is getting many answers that have little basis in fact. The Metallic Metals Act illustration may be challenged as unusual, but in another case a Gallup report revealed that 45 percent of the persons surveyed did not know what a "lobbyist in Washington" was and 88 percent could not give
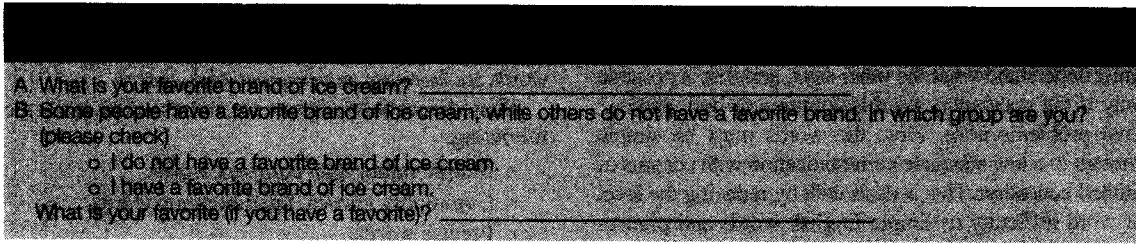
a correct description of "jurisdictional strike."[2] This points to the need for operational definitions as part of question wording.

Presumed Knowledge The question designer should consider the participants' information level when determining the content and appropriateness of a question. In some studies, the degree of participant expertise can be substantial, and simplified explanations are inappropriate and discourage participation. In asking the public about gross margins in menswear stores, we would want to be sure the "general-public" participant understands the nature of "gross margin." If our sample unit were a merchant, explanations might not be needed. A high level of knowledge among our sample units, however, may not eliminate the need for operational definitions. Among merchants, gross margin per unit in dollars is commonly accepted as the difference between cost and selling price; but when offered as a percentage rather than a dollar figure, it can be calculated as a percentage of unit selling price or as a percentage of unit cost. A participant answering from the "cost" frame of reference would calculate gross margin at 100 percent; another participant, using the same dollars and the "selling price" frame of reference, would calculate gross margin at 50 percent. If a construct is involved and differing interpretations of a concept are feasible, operational definitions may still be needed.

Recall and Memory Decay The adequacy problem also occurs when you ask questions that overtax participants' recall ability. People cannot recall much that has happened in their past, unless it was dramatic. Your mother may remember everything about your arrival if you were her first child: the weather, time of day, even what she ate prior to your birth. If you have several siblings, her memory of subsequent births may be less complete. If the events surveyed are of incidental interest to participants, they will probably be unable to recall them correctly even a short time later. An unaided recall question, "What radio programs did you listen to last night?" might identify as few as 10 percent of those individuals who actually listened to a program.[3]

Balance (General versus Specific) Answering adequacy also depends on the proper balance between generality and specificity. We often ask questions in terms too general and detached from participants' experiences. Asking for average annual consumption of a product may make an unrealistic demand for generalization on people who do not think in such terms. Why not ask how often the product was used last week or last month? Too often participants are asked to recall individual use experiences over an extended time and to average them for us. This is asking participants to do the researcher's work and encourages substantial response errors. It may also contribute to a higher refusal rate and higher discontinuation rate.

> **Exhibit 14a-1** A Test of Alternative Response Strategies

A. What is your favorite brand of ice cream?
B. Some people have a favorite brand of ice cream, while others do not have a favorite brand. In which group are you? (please check)
    o I do not have a favorite brand of ice cream.
    o I have a favorite brand of ice cream.
    What is your favorite (if you have a favorite)? _____

> **Exhibit 14a-2** Results of Alternative Response Strategies Test

Response	Version A	Version B
Named a favorite brand	77%*	39%*
Named a favorite flavor rather than a brand	19	18
Had no favorite brand	4	43
Total	100%	100%
	n = 57	n = 56

*Significant difference at the 0.001 level.

There is a danger in being too narrow in the time frame applied to behavior questions. We may ask about movie attendance for the last seven days, although this is too short a time span on which to base attendance estimates. It may be better to ask about attendance, say, for the last 30 days. There are no firm rules about this generality-specificity problem. Developing the right level of generality depends on the subject, industry, setting, and experience of the question designer. **Objectivity** The ability of participants to answer adequately is also often distorted by questions whose content is biased by what is included or omitted. The question may explicitly mention only the positive or negative aspects of the topic or make unwarranted assumptions about the participant's position. Consider Exhibit 14a-1, an experiment in which two forms of a question were asked. Fifty-seven randomly chosen graduate business students answered version A, and 56 answered version B. Their responses are shown in Exhibit 14a-2. The probable cause of the difference in level of brand preference expressed is that A is an unsupported assumption. It assumes and suggests that everyone has a favorite brand of ice cream and will report it. Version B indicates the participant need not have a favorite.

A deficiency in both versions is that about one participant in five misinterpreted the meaning of the term *brand*. This misinterpretation cannot be attributed to low education,

*Word confusion difficulties are discussed later in this appendix.

low intelligence, lack of exposure to the topic, or quick or lazy reading of the question. The subjects were students who had taken at least one course in marketing in which branding was prominently treated.*

Will the Participants Answer Willingly?

Sensitive Information Even if participants have the information, they may be unwilling to give it. Some topics are considered too sensitive to discuss with strangers. These vary from person to person, but one study suggests the most sensitive topics concern money matters and family life.[4] More than one-fourth of those interviewed mentioned these as the topics about which they would be "least willing to answer questions." Participants of lower socioeconomic status also included political matters in this "least willing" list.

Participants also may be unwilling to give correct answers for ego reasons. Many exaggerate their incomes, the number of cars they own, their social status, and the amount of high-prestige literature they read. They also minimize their age and the amount of low-prestige literature they read. Many participants are reluctant to try to give an adequate response. Often this will occur when they see the topic as irrelevant to their own interests or to their perception of the survey's purpose. They participate halfheartedly, often answer with "don't know," give negative replies, give stereotypical responses, or refuse to be interviewed.

You can learn more about crafting questions dealing with sensitive information by reading the CD Closeup "Measuring Attitudes on Sensitive Subjects."

Question Wording

Shared Vocabulary Because surveying is an exchange of ideas between interviewer and participant, each must understand what the other says, and this is possible only if the vocabulary used is common to both parties.[5] Two problems arise. First, the words must be simple enough to allow adequate communication with persons of limited education. This is dealt with by reducing the level of word difficulty to simple English words and phrases (more is said about this in the section on word clarity).

Technical language is the second issue. Even highly educated participants cannot answer questions stated in unfamiliar technical terms. Technical language also poses difficulties for interviewers. In one study of how corporation executives handled various financial problems, interviewers had to be conversant with technical financial terms. This necessity presented the researcher with two alternatives—hiring people knowledgeable in finance and teaching them interviewing skills or teaching financial concepts to experienced interviewers.[6] This vocabulary problem also exists where similar or identical studies are conducted in different countries and multiple languages.

A great obstacle to effective question wording is the choice of words. Questions to be asked of the public should be restricted to the 2,000 most common words in the English language.[7] Even the use of simple words is not enough. Many words have vague references or meanings that must be gleaned from their context. In a repair study, technicians were asked, "How many radio sets did you repair last month?" This question may seem unambiguous, but participants interpreted it in two ways. Some viewed it as a question of them alone; others interpreted "you" more inclusively, as referring to the total output of the shop. There is also the possibility of misinterpreting "last month," depending on the timing of the questioning. Using "during the last 30 days" would be much more precise and unambiguous. Typical of the many problem words are these: *any, could, would, should, fair, near, often, average,* and *regular.* One author recommends that after stating a question as precisely as possible, we should test each word against this checklist:

- Does the word chosen mean what we intend?
- Does the word have multiple meanings? If so, does the context make the intended meaning clear?
- Does the word chosen have more than one pronunciation? Is there any word with similar pronunciation with which the chosen word might be confused?
- Is a simpler word or phrase suggested or possible?[8]

We cause other problems when we use abstract concepts that have many overtones or emotional qualifications.[9] Without concrete referents, meanings are too vague for the researcher's needs. Examples of such words are *business, government,* and *society.*

Shared vocabulary issues are addressed by using the following:

- Simple rather than complex words.
- Commonly known, unambiguous words.
- Precise words.
- Interviewers with content knowledge.

Unsupported Assumptions Unwarranted assumptions contribute to many problems of question wording. A metropolitan newspaper, *Midwest Daily,* conducted a study in an attempt to discover what readers would like in its redesigned lifestyle section. One notable question asked readers: "Who selects your clothes? You or the man in your life?" In this age of educated, working, independent women, the question managed to offend a significant portion of the female readership. In addition, *Midwest Daily* discovered that many of its female readers were younger than researchers originally assumed and the only man in their lives was their father, not the spousal or romantic relationship alluded to by the questions that followed. Once men reached this question, they assumed that the paper was interested in serving only the needs of female readers. The unwarranted assumptions built into the questionnaire caused a significantly smaller response rate than expected and caused several of the answers to be uninterpretable.

Frame of Reference Inherent in word meaning problems is also the matter of a frame of reference. Each of us understands concepts, words, and expressions in light of our own experience. The U.S. Bureau of the Census wanted to know how many people were in the labor market. To learn whether a person was employed, it asked, "Did you do any work for pay or profit last week?" The researchers erroneously assumed there would be a common frame of reference between the interviewer and participants on the meaning of *work.* Unfortunately, many persons viewed themselves primarily or foremost as homemakers or students. They failed to report that they also worked at a job during the week. This difference in frame of reference resulted in a consistent underestimation of the number of people working in the United States.

In a subsequent version of the study, this question was replaced by two questions, the first of which sought a statement on the participant's major activity during the week. If the participant gave a nonwork classification, a second question was asked to determine if he or she had done any work for pay besides this major activity. This revision increased the estimate of total employment by more than 1 million people, half of them working 35 hours or more per week.[10]

> **Exhibit 14a-3** Split Test of Alternative Question Wording

Source: Hadley Cantril, ed., Gauging Public Opinion (Princeton, NJ: Princeton University Press, 1944), p. 48.

The frame of reference can be controlled in two ways. First, the interviewer may seek to learn the frame of reference used by the participant. When asking participants to evaluate their reasons for judging a retail store as unattractive, the interviewer must learn the frames of reference they use. Is the store being evaluated in terms of its particular features and layout, the failure of management to respond to a complaint made by the participant, the preference of the participant for another store, or the participant's recent difficulty in returning an unwanted item?

Second, it is useful to specify the frame of reference for the participant. In asking for an opinion about the new store design, the interviewer might specify that the question should be answered based on the participant's opinion of the layout, the clarity and placement of signage, the ease of finding merchandise, or another frame of reference.

Biased Wording Bias is the distortion of responses in one direction. It can result from many of the problems already discussed, but word choice is often the major source. Obviously such words or phrases as *politically correct* or *fundamentalist* must be used with great care. Strong adjectives can be particularly distorting. One alleged opinion survey concerned with the subject of preparation for death included the following question: "Do you think that decent, low-cost funerals are sensible?" Who could be against anything that is *decent* or *sensible?* There is a question about whether this was a legitimate survey or a burial service sales campaign, but it shows how suggestive an adjective can be.

Congressional representatives have been known to use surveys as a means of communicating with their constituencies. Questions are worded, however, to imply the issue stance that the representative favors. Can you tell the representative's stance in the following question?

Example: Would you have me vote for a balanced budget if it means higher costs for the supplemental Social Security benefits which you have already earned?

We can also strongly bias the participant by using prestigious names in a question. In a historic survey on whether the war and navy departments should be combined into a single defense department, one survey said, "General Eisenhower says the army and navy should be combined," while the other version omitted his name. Given the first version (name included), 49 percent of the participants approved of having one department; given the second version, only 29 percent favored one department.[11] Just imagine using Michael Jordan's or Shaq O'Neill's name in a survey question asked of teen boys interested in basketball. The power of aspirational reference groups to sway opinion and attitude is well established in advertising; it shouldn't be underestimated in survey design.

We also can bias response through the use of superlatives, slang expressions, and fad words. These are best excluded unless they are critical to the objective of the question. Ethnic references should also be stated with care.

Personalization How personalized should a question be? Should we ask, "What would you do about . . . ?" Or should we ask, "What would people with whom you work do about . . . ?" The effect of personalization is shown in a classic example reported by Cantril.[12] A split test—where a portion of the sample received one question, with another portion receiving a second question—was made of a question concerning attitudes about the expansion of U.S. armed forces in 1940, as noted in Exhibit 14a-3.

These and other examples show that personalizing questions changes responses, but the direction of the influence is not clear. We cannot tell whether personalization or no personalization is superior. Perhaps the best that can be said is that when either form is acceptable, we should choose that which appears to present the issues more realistically. If there are doubts, then split survey versions should be used (one segment of the sample should get one question version, while a second segment should receive the alternative question version).

Adequate Alternatives Have we adequately expressed the alternatives with respect to the purpose of the question? It is usually wise to express each alternative explicitly to avoid bias. This is illustrated well with a pair of questions that were asked of matched samples of participants.[13] The question forms that were used are noted in Exhibit 14a-4.

> **Exhibit 14a-4** Expressing Alternatives
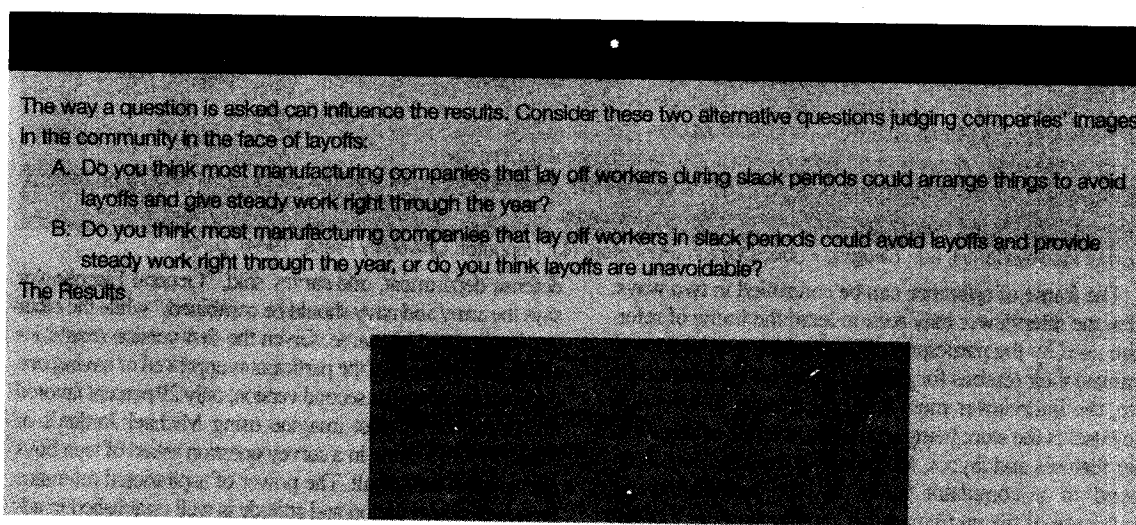
Source: Hadley Cantril, ed., *Gauging Public Opinion* (Princeton, NJ: Princeton University Press, 1944), p. 48.

Often the above issues are simultaneously present in a single question. Exhibit 14a-5 reveals several questions drawn from actual mail surveys. We've identified the problem issues and suggest one solution for improvement. While the suggested improvement might not be the only possible solution, it does correct the issues identified. What other solutions could be applied to correct the problems identified?

Response Strategy

The objectives of the study; characteristics of participants, especially their level of information, level of motivation to participate, and ease of communication; the nature of the topic(s) being studied; the type of scale needed; and your analysis plan dictate the response strategy. Examples of the strategies described in Chapter 14 and discussed in detail in Chapters 12 and 13 are found in Exhibit 14-6.

Objective of the Study   If the objective of the question is only to classify the participant on some stated point of view, then the closed question will serve well. Assume you are interested only in whether a participant approves or disapproves of a certain corporate policy. A closed question will provide this answer. This response strategy ignores the full scope of the participant's opinion and the events that helped shape the attitude at its foundation. If the objective is to explore this wider territory, then an open-ended question (free-response strategy) is preferable.

Open-ended questions are appropriate when the objective is to discover opinions and degrees of knowledge. They are also appropriate when the interviewer seeks sources of information, dates of events, and suggestions or when probes are used to secure more information. When the topic of a question is outside the

participant's experience, the open-ended question may offer the better way to learn his or her level of information. Closed questions are better when there is a clear frame of reference, the participant's level of information is predictable, and the researcher believes the participant understands the topic.

Open-ended questions also help to uncover certainty of feelings and expressions of intensity, although well-designed closed questions can do the same.

Thoroughness of Prior Thought   If a participant has developed a clear opinion on the topic, a closed question does well. If an answer has not been thought out, an open-ended question may give the participant a chance to ponder a reply, and then elaborate on and revise it.

Communication Skill   Open-ended questions require a stronger grasp of vocabulary and a greater ability to frame responses than do closed questions.

Participant Motivation   Experience has shown that closed questions typically require less motivation and answering them is less threatening to participants. But the response alternatives sometimes suggest which answer is appropriate; for this reason, closed questions may be biased.

While the open-ended question offers many advantages, closed questions are generally preferable in large surveys. They reduce the variability of response, make fewer demands on interviewer skills, are less costly to administer, and are much easier to code and analyze. After adequate exploration and testing, we can often develop closed questions that will perform as effectively as open-ended questions in many situations. Experimental studies suggest that closed questions are equal or superior to open-ended questions in many more applications than is commonly believed.[14]

> **Exhibit 14a-5** Reconstructing Questions

Problem/Solution	Poor Measurement Question	Improved Measurement Question

Problems: Checklist appears to offer options that are neither exhaustive nor mutually exclusive. Also, it doesn't fully address the content needs of understanding why people choose a hotel when they travel for personal reasons versus business reasons. **Solution:** Organize the alternatives. Create subsets within choices; use color or shading to highlight subsets. For coding ease, expand the alternatives so the participant does not frequently choose "Other."	If your purpose for THIS hotel stay included personal pleasure, for what ONE purpose specifically?
	 ❏ Visit friend/relative
	 ❏ Weekend escape
	 ❏ Sporting event
	 ❏ Sightseeing
	 ❏ Family event
	 ❏ Vacation
	 ❏ Other _____
		Which reason BEST explains your purpose for THIS personal pleasure hotel stay?
		 ❏ Dining
		 ❏ Shopping
		 ❏ Entertainment
		    . . . was this for a . . . ❏ Sport-related event?
		                  ❏ Theater, musical, or other performance?
		                  ❏ Museum or exhibit?
		 ❏ Visit friend/relative
		    . . . was this for a special event? ❏ YES ❏ NO
		 ❏ Vacation
		    . . . was this primarily for . . . ❏ Sightseeing?
		                  ❏ Weekend escape?
		 ❏ Other: _____

Problem: Nonspecific time frame; likely to experience memory decay; nonspecific screen (not asking what you really need to know to qualify a participant). **Solution:** Replace "ever" with a more appropriate time frame; screen for the desired behavior.	Have you ever attended a college basketball game? ❏ YES ❏ NO	In the last six months, have you been a spectator at a basketball game played by college teams on a college campus? ❏ YES ❏ NO

>appendix14b

Pretesting Options and Discoveries

Pretesting is a critical activity for successful development of a survey. We explore here the purposes and methods for effectively pretesting questions and instruments.

Pretesting Options

There are various types of pretesting that can be used to refine an instrument. They range from obtaining informal reviews by colleagues to creating conditions similar to those of the final study.

Researcher Pretesting

Designers typically test informally in the initial stages and build more structure into the tests along the way. Fellow instrument designers can do the first-level pretest. One way to accomplish this is to have researchers divided into teams. One team writes the survey, while the other critically reviews it. The reviewers' and researchers' many differences of opinion are likely to create numerous suggestions for improvement. Usually at least two or three drafts can be effectively developed by bringing research colleagues into the process.

Participant Pretesting

Participant pretests require that the questionnaire be field-tested by sample participants or participant surrogates (individuals with characteristics and backgrounds similar to those of the desired participants).

Field pretests involve distributing the test instrument exactly as the actual instrument will be distributed. Most studies use two or more pretests. National projects may use one trial to examine local reaction and another to check for regional differences. Although many researchers try to keep pretest conditions and times close to what they expect for the actual study, personal interview and telephone limitations make it desirable to test in the evenings or on weekends in order to interview people who are not available for contact at other times.

Test mailings are useful, but it is often faster to use a substitute procedure. In the MindWriter example, the managers who were interviewed in the exploratory study were later asked to review the pilot questionnaire. The interviewers left them alone and returned later. Upon their return, they went over the questions with each manager. They explained that they wanted the manager's reactions to question clarity and ease of answering. After several such interviews, the instrument was revised and the testing process was repeated with customers. With minor revision, the questionnaire was reproduced and prepared for insertion into the computer packing material.

Collaborative Pretests

Different approaches taken by interviewers and the participants' awareness of those approaches affect the pretest. If the researcher alerts participants to their involvement in a preliminary test of the questionnaire, the participants are essentially being enlisted as collaborators in the refinement process. Under these conditions, detailed probing of the parts of the question, including phrases and words, is appropriate. Because of the time required for probing and discussion, it is likely that only the most critical questions will be reviewed. The participant group may therefore need to be conscripted from colleagues and friends to secure the additional time and motivation needed to cover an entire questionnaire. If friends or associates are used, experience suggests that they introduce more bias than strangers, argue more about wording, and generally make it more difficult to accomplish other goals of pretesting such as timing the length of questions or sections.[1]

Occasionally, a highly experienced researcher may improvise questions during a pretest. When this occurs, it is essential to record the interview or take detailed notes so that the questionnaire may be reconstructed later. Ultimately, a team of interviewers would be required to follow the interview schedule's prearranged sequence of questions. Only experienced investigators should be free to depart from the interview schedule during a pretest and explore participants' answers by adding probes.

Noncollaborative Pretests

When the researcher does not inform the participant that the activity is a pretest, it is still possible to probe for reactions but without the cooperation and commitment of time provided by collaborators. The comprehensiveness of the effort also suffers because of flagging cooperation. The virtue of this approach is that the questionnaire can be tested under conditions approaching those of the final study. This realism is similarly useful for training interviewers.

Pretesting Discoveries

Participant Interest

An important purpose of pretesting is to discover participants' reactions to the questions. If participants do not find the experience stimulating when an interviewer is physically present, how will they react on the phone or in the self-administered mode? Pretesting helps discover where repetitiveness or redundancy is bothersome or what topics were not covered that the participant expected. An alert interviewer will look for questions or groups of questions that the participant perceives to be sensitive or threatening or topics about which the participant knows nothing.

Meaning

Questions that we borrow or adapt from the work of others carry an authoritativeness that may prompt us to avoid pretesting them, but they are often most in need of examination. Are they still timely? Is the language relevant? Do they need context from adjacent questions? Newly constructed questions should be similarly checked for meaningfulness to the participant. Does the question evoke the same meaning as that intended by the researcher? How different is the researcher's frame of reference from that of the average participant? Words and phrases that trigger a "what do you mean?" response from the participant need to be singled out for further refinement.

Question Transformation

Participants do not necessarily process every word in the question. They also may not share the same definitions for the terms they hear. When this happens, participants modify the question to make it fit their own frame of reference or simply change it so that it makes sense to them. Probing is necessary to discover how participants have transformed the question when this is suspected.[2]

Continuity and Flow

In self-administered questionnaires, questions should read effortlessly and flow from one to another and from one section to another. In personal and telephone interviews, the sound of the question and its transition must be fluid as well. A long set of questions with 9-point scales that worked well in a mail instrument would not be effective on the telephone unless you were to ask participants to visualize the scale as the touch keys on their phone. Moreover, transitions that may appear redundant in a self-administered questionnaire may be exactly what needs to be heard in personal or telephone interviewing.

Question Sequence

Question arrangement can play a significant role in the success of the instrument. Research authorities recommend starting with stimulating questions and placing sensitive questions last. Since questions concerning income and family life are most likely to be refused, this is good advice for building trust before getting to classification questions that might lead to a refusal situation. However, interest-building questions need to be tested first to be sure they are stimulating. Pretesting with a large enough sample of participants permits some experimentation with question sequence.

Skip Instructions

In interviews and questionnaires, skip patterns and their contingency sequences may not work as envisioned on paper. Skip instructions are designed to route or sequence the participant to another question contingent on his or her answer to the previous question *(branched questions)*. Pretesting in the field helps to identify problems with skip instructions or symbols (e.g., box-and-arrow schematic) that the designers may not have thought of. By correcting these instructions in the revision stage, we also avoid problems with flow and continuity.

Variability

Making sure that question alternatives cover the range of possible participant answers is an important purpose of pretesting. With 25 to 100 participants in the pretest sample, statistical data on the proportion of participants answering yes or no or marking "strongly agree" to "strongly disagree" can supplement the qualitative assessment provided by the pretest interviewers. This information is useful for sample size calculations and for getting preliminary indications of reliability problems with scaled questions. When using a very small pretest sample of participants, pretesting cannot provide definitive quantitative conclusions. Small samples can, however, deliver an early warning about survey questions that may not discriminate among participants or can identify sections of the survey where meaningful subgrouping may occur in the final sample.

Length and Timing

Most draft questionnaires or interview schedules suffer from lengthiness. By timing each question and section, the researcher is in a better position to make decisions about modifying or cutting material. In personal and telephone interviews, labor is a project expense. Thus, if the budget influences the final length of the questionnaire, an accurate estimate of elapsed time is essential. Videotaped or audiotaped pretests may also be used for this purpose. Their function in reducing errors in data recording is widely accepted.

When Surveying Doesn't Work

Sometimes surveying will not secure the information needed. A classic example concerns a survey conducted to discover magazines read by participants. An unusually high rate was reported for prestigious magazines, and an unusually low rate was reported for tabloid magazines. The study was revised so that the subjects, instead of being interviewed, were asked to contribute their old magazines to a charity drive (an observation study). The collection gave a more realistic estimate of readership of both types of magazines.[3]

Most researchers have found that the survey is a very powerful tool in their research methods arsenal. It is only a matter of careful attention to detail and practice that will have you joining their ranks.